

DOCUMENT RESUME

ED 111 739

SO 008 592

AUTHOR Contreras, Gloria
TITLE Mastery Learning: The Relation of Different Criterion Levels and Aptitude to Achievement, Retention, and Attitude in a Seventh Grade Geography Unit.
PUB DATE 75
NOTE 230p.; Ed. D. Dissertation, University of Georgia
AVAILABLE FROM Geography Curriculum Project, 107 Dudley Hall, University of Georgia, Athens, Georgia 30601 (\$5.00)
EDRS PRICE MF-\$0.76 HC-\$12.05 Plus Postage
DESCRIPTORS *Academic Achievement; Achievement Rating; Doctoral Theses; Educational Research; Geography; *Geography Instruction; Knowledge Level; *Learning; Low Ability Students; *Retention Studies; Secondary Education; *Student Attitudes; Student Improvement
IDENTIFIERS Geography Curriculum Project; Mastery Learning

ABSTRACT

This study is an assessment of the effect of three criterion mastery levels and aptitude on the achievement, retention, and attitude of seventh grade students using a population geography text. The three criterion levels used were 90 percent, 80 percent, and 70 percent of the correct responses on each of 41 lessons in "Population Growth in the United States and Mexico." Aptitude was measured by a word meaning text. Aptitude was a major independent variable because of the premise that mastery procedures may overcome achievement difficulties of low aptitude students. Contrary to most of the literature on mastery learning, the results indicate that mastery or formative exercises did not contribute to higher levels of achievement on the summative test. The study also failed to show that achievement, retention, or attitude toward the unit differed among the three criterion groups. The study confirmed that previous knowledge, as measured by a word meaning test, was a more potent factor in achievement than were differential criterion levels.
(Author/DE)

* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *

AUG 12 1975

MASTERY LEARNING: THE RELATION OF DIFFERENT
CRITERION LEVELS AND APTITUDE TO ACHIEVEMENT,
RETENTION, AND ATTITUDE IN A SEVENTH GRADE GEOGRAPHY UNIT

by

GLORIA CONTRERAS

B.S., The University of Texas at El Paso, 1969

M.Ed., The University of Texas at El Paso, 1973

A Dissertation Submitted to the Graduate Faculty
of the University of Georgia in Partial Fulfillment
of the

Requirements for the Degree

DOCTOR OF EDUCATION

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRE-
SENT OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

ATHENS, GEORGIA

1975

PERMISSION TO REPRODUCE THIS COPY-
RIGHTED MATERIAL HAS BEEN GRANTED BY

M. J. Rice

TO ERIC AND ORGANIZATIONS OPERATING
UNDER AGREEMENTS WITH THE NATIONAL IN-
STITUTE OF EDUCATION. FURTHER REPRO-
DUCTION OUTSIDE THE ERIC SYSTEM RE-
QUIRES PERMISSION OF THE COPYRIGHT
OWNER

ACKNOWLEDGEMENTS

My sincere thanks go to my professors at the University of Georgia for the training they have given me. A special thanks is extended to Dr. M. J. Rice, my major professor and Director of the Anthropology and Geography Curriculum Projects, for his encouragement and guidance.

Appreciation is also extended to the following personnel for helping make this study possible:

El Paso Independent School District,

Superintendent: Mr. J. M. Whitaker

Director of Research: Dr. Guy McNiel

Social Studies Consultants: Mr. Clinton Hartman

Mrs. Yolanda Rey

Mr. Shelby Martin

Canyon Hills School:

Principal, Mrs. Marie Pedraza,

Teacher, Mrs. Connie Pope

Crosby School:

Principal, Mr. Lamar A. Taylor

Teacher, Mrs. Carol Aune

Guillen School:

Principal, Mr. George Chriss

Teacher, Mrs. Lynne E. Payne

Henderson School:

Principal, Mr. Joe Ontiveros

Teacher, Mr. Philip Duncan

White School:

Principal, Mr. Charles Holmes

Teacher, Mr. Benjamin F. Gomersall

FOREWORD

This study was undertaken as part of the continuing research and development of the Geography Curriculum Project, University of Georgia.

The content focus of the Geography Curriculum Project is the preparation of supplementary units for the elementary grades, emphasizing the organizing concepts of the discipline of geography. The research focus is the testing of some psychological construct of learning, such as the nature of concepts, Ausubel's reception learning model, Bloom's mastery learning, or Bruner's discovery hypothesis, under normal conditions of school instruction.

The Geography Curriculum Project thus serves as a small research and development center. It develops new materials and measurement instruments, field tests and evaluates materials, and facilitates the training of doctoral students in geographic education.

The Geography Curriculum Project was initiated as a result of a study of geographic content in elementary social science texts, manuals, and study guides. The evidence indicated that elementary geography is primarily presented as a discrete body of facts, with little attention to the organizing concepts of geography which help to analyze, interpret and integrate physical and cultural phenomena. The development of systematic geography units helps to clarify

the teaching of geographic knowledge and concepts. The research emphasis answers questions relating to the structuring of materials and their use in teaching geography.

TABLE OF CONTENTS

CHAPTER	PAGE
I. PURPOSE, BACKGROUND AND HYPOTHESES OF THE STUDY	1
Background	2
Objectives and Research Hypotheses	12
Definition of Terms	18
II. REVIEW OF THE LITERATURE	20
The History of Mastery Learning	20
Bloom and Keller Mastery Learning Strategies	31
Mastery Learning Research	35
Achievement	36
Retention	42
Attitude	46
Time	49
Aptitude and Achievement Under Conditions of Mastery and Non- Mastery	53
Mastery Learning and the Social Studies	58
Origins of the Present Study	62
Description of the Block Study	63

CHAPTER	PAGE
Results and Discussion	64
III. PROCEDURES AND METHODOLOGIES	71
Description of Treatment	71
Treatment Variable	71
Instructional Unit	72
Sample Selection	77
Duration of the Study	77
Pilot Study	78
Orientation of Teachers	81
Experimental Study	82
Description of Data Collection	83
Word Meaning Test	83
Achievement Test	85
Attitude Scale	85
Description of Contextual Variables	88
Community and the School District	88
Characteristics of Schools in the Study	90
Characteristics of Teachers in the Study	94
Summary of Contextual Variables	97
Experimental Design	99
Random Assignment of Classes to Treatment Groups	99
Distribution of Students by Treat- ment and Vocabulary Level	104

CHAPTER

PAGE

Vocabulary Scores, Grade Equivalent, and National Percentile Rank	108
Unit of Statistical Analysis	114
Statistical Procedures	114
Statement of the Statistical Hypotheses	115
Significance Level	116
Limitations of the Study	117
IV. RESULTS AND DISCUSSION OF THE FINDINGS	120
Presentation of the Findings	120
Discussion of the Findings	133
V. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS FOR FOLLOW-UP RESEARCH	145
Procedures	146
Findings	147
Conclusions	148
Recommendations for Follow-up Research	151
REFERENCES	155
APPENDICES	166
A Table of Contents and Three Sample Lessons from <u>Population Growth in the United States and Mexico</u>	167

APPENDICES

PAGE

B	Pupils' Personal Score Logs for the 90%, 80%, and 70% Criterion Levels	177
C	Treatment Directions to the Student and Teacher	181
D	Questions for Developing Map and Graph Reading Skills	184
E	Geography Achievement Posttest	190
F	<u>Attitude Toward Any Subject Scale</u> , Short A Form	200
G	Bibliography of Mastery Learning Studies at the College Level	203
H	Request for Authorization to Conduct Re- search in the El Paso Public Schools	208
I	Time Table for Pilot and Experimental Study	215

LIST OF TABLES

TABLE		PAGE
2.1	Achievement Under Conditions of Mastery and Non-Mastery Learning	37
2.2	Retention Under Conditions of Mastery and Non-Mastery Learning	44
2.3	Attitudinal Outcomes Under Conditions of Mastery and Non-Mastery Learning	47
2.4	Learning Time Under Conditions of Mastery and Non-Mastery Learning	51
2.5	Aptitude in Relation to Mastery Learning	54
2.6	Social Studies and Mastery Learning Research	59
2.7	Variables Used by Block (1970) and Contreras (1975)	67
3.1	Computed Reading Ease Scores According to the Rudolf Flesch Readability Formula	74
3.2	Computed Reading Ease Score According to Fry Readability Formula	76
3.3	School Characteristics Summary Table	91
3.4	Teacher Characteristics Summary Table	95
3.5	Vocabulary Levels Across Two Types of Schools	100

TABLE		PAGE
3.6	Distribution of Classes Across Treatment by Ethnic Groups	101
3.7	Experimental Layout (ANOVA)	102
3.8	Mean Vocabulary Achievement Scores by Treatment Groups	105
3.9	Number of Students by Vocabulary Level, Class, and Treatment	106
3.10	Mean Reading Scores and Grade Equivalents by Treatment, Class, and Vocabulary Level	109
3.11	Mean Achievement Scores by Treatment, Class, and Vocabulary Level	110
3.12	Mean Retention Scores by Treatment, Class and Vocabulary Level	111
3.13	Mean Attitude Toward the Unit Scores by Treatment, Class, and Vocabulary Level	112
3.14	Mean Reading Scores, Grade Equivalent, and National Percentile Rank for Vocabulary Groups	113
4.1	Analysis of Variance for Treatment, Vo- cabulary Level, and Interaction - Achievement Posttest	122
4.2	Achievement Mean Scores by Treatment and Vocabulary Level	123

TABLE

PAGE

4.3	Analysis of Variance for Treatment, Vocabulary Level, and Interaction - Achievement Delayed Posttest	126
4.4	Retention Mean Scores by Treatment and Vocabulary Level	127
4.5	Analysis of Variance for Treatment, Vocabulary Level, and Interaction - Attitude Posttest	129
4.6	Attitude Mean Scores by Treatment and Vocabulary Level	130
4.7	Summary of Univariate Tests of Signifi- cance: Interaction and Main Effects	132
4.8	Number of Students Achieving Criterion on the Summative Test by Treatment and Vocabulary Level and Mean Achievement, Retention, and Attitude, by Treatment	135
4.9	Number and Percentage of Students Not Com- pleting the Unit and Mean Number of Lessons Completed by Treatment and Vocabulary Level	140

CHAPTER I

PURPOSE, BACKGROUND AND HYPOTHESES OF THE STUDY

This study is an assessment of the effect of three criterion mastery levels and aptitude on the achievement, retention, and attitude of seventh grade students using a population geography text. The three criterion levels used in this study were 90%, 80%, and 70% of the correct responses on each of 41 lessons in Population Growth in the United States and Mexico. Aptitude was measured by a word meaning test. Aptitude was a major independent variable of the study because a basic premise of mastery learning is that mastery procedures may overcome achievement difficulties of low aptitude students.

This study, like other mastery learning treatments in social studies using anthropology (Gaines, 1971) and geography materials (Jones, 1974; Fagan, 1975; Myers, 1975), did not find that mastery on formative exercises contributed to higher levels of achievement on the summative test, a finding contrary to most of the literature on mastery learning, as described in the review of literature. The study also failed to show that achievement, retention, or attitude toward the unit differed among the three criterion groups, a finding

specifically contrary to that reported by Block (1970) with a four-day mathematics experiment. The study confirmed that previous knowledge, as measured by a word meaning test, was a more potent factor in achievement than were differential criterion levels.

Background

Central to the history of formal education has been the problem of how to approach instruction systematically so that most students can learn what schools assign. Efforts to improve student school performance have been noted since the formulation of St. Ignatius Loyola's Ratio Studiorum (Fitzpatrick, 1933) to present day mastery learning processes (Bloom, 1968). Yet despite advances in educational psychology and the investment of immense amounts of money, time, and effort, every year a large proportion of students achieve at low levels of performance or altogether fail to succeed in school tasks. Low performance is cumulative, so these students increasingly fall behind in their school performance. The criticism of such educational shortcomings, however, goes on in an atmosphere of frustration and futility because there is no agreement on how correction is to be applied (Findley, 1972).

There is no general plan or correction because educators disagree considerably on what the central purpose of formal education should be (Bigge, 1972; Biehler, 1971). Some maintain that it is the intellectual development of the

child or the cultivation of his cognitive competence' (Bagley, 1938; Educational Policies Commission, 1961; Ebel, 1972).

Others give priority to a child's personal and social development, his feelings of self worth and happiness, and his sense of belonging and security (Wilhelms, 1967).

Those who promote the intellectual approach to education tend to support conventional school systems and practices and to seek educational reforms through successive improvements in traditional materials and instructional methods. Those who give priority to a pupil's feelings and adjustment tend to favor flexibility and freedom for pupils, teachers, curriculum, and procedures. Many advocates of this latter orientation favor getting rid of textbooks, tests, grades, and conventional classroom arrangements (Ebel, 1972).

A discussion of these two extreme approaches to education, from a framework of systematic analysis, is given by Smith, Stanley, and Shores (1957) under the labels subject matter and activity curriculum. Many aspects of the activity curriculum have been revived in recent years under such broad terms as the humanistic school (Kruger, 1974) and the open school (Bercheck & Tauss, 1973; Haynes, 1973). In 1970, Barth and Shermis attempted to relate these divergent trends specifically to social studies under the three categories of citizenship transmission, social science, and reflective inquiry.

Within the two orientations, intellectual or child centered, there is considerable diversity among the advocates

of either outlook. Whether the orientations be termed intellectual or humanist, conservative or progressive, scholarly or reformist, structured or unstructured, or formalistic and naturalistic, there are important differences between points of view about the educational process with important consequences in the choice of methodology and in results (Jarvis & Rice, 1972).

To these differences in philosophic premises about the nature of knowledge and psychological assumptions about the process of learning may be attributed many of the differences in educational theory and practice and the often inconsistent application of reform efforts. In the current scene, for example, advocates of competency based education concurrently accept the principles of Thorndikian (1906) and Skinnerian (1954) scientific educational management as indicated by behavioral objectives and operant conditioning, while espousing a greater amount of freedom for the child and teacher in developing an emergent curriculum.

These contrasting methodologies have been described by Jarvis and Rice (1972) as formalism and naturalism. A formal approach to education emphasizes the transmission of subject matter and efficient teaching methodologies controlled by the teacher. The natural approach emphasizes the growth of the child and learning methods emphasizing child activity. Formalism assumes that rationality is based in knowledge, while naturalism places more emphasis on emotions and sentiments and individual expression.

In the nineteenth century, formal methods dominated the American educational scene, despite child-centered approaches adopted from Pestalozzian concepts (Jarvis & Rice, 1972).

The Lancastrian system, in particular, was immensely popular in the early part of the century, combining elements of economic efficiency with clearly defined and sequenced learning procedures (Cubberly, 1934). In the twentieth century, formalism has prevailed in school practice while naturalistic methods have dominated pedagogic theory. However, the Progressive Movement embraced both formalistic and naturalistic approaches. Formalistic approaches included connectionist psychology (Thorndike, 1924, 1932) and self-tutorial programs (Washburne, 1926, 1940). The most popular legacy of Progressivism, however, was naturalistic and included the project method (Kilpatrick, 1926) and the child-centered school (Rugg and Shumaker, 1928).

Formalism is most frequently associated with an elitist educational view. Of contemporary measurement psychology and theorists in the United States, Ebel (1972) is one of the most articulate spokesman for a highly intellectual approach to education, with schools serving as a screening mechanism for social purposes. From his perspective, a good school will not try to banish failure. He affirms that "Whenever goals are reasonably challenging and standards reasonably high, failure are bound to occur. Kept to a minimum they constitute a valuable part of a person's education" (1972, p. 48). He suggests that one reason so few achieve

excellence is an over-commitment to a uniform program of general or liberal education for all.

In contrast to this elitist approach to education, there is the idea that schools should exist without failure (Carroll, 1963; Bloom, 1968). Proponents of this approach more readily identify themselves with naturalistic approaches. Some advocates of formal approaches use educational technology as a means of eliminating failure and bringing high performance levels to all students, notwithstanding differences in individual aptitude. Basically, this is one premise of compensatory education, that diagnostic and prescriptive instruction management systems will reduce, if not completely eliminate, the learning difficulties of children, traditionally attributed to differences in aptitude.

Many compensatory educational efforts have been directed to environmental and social factors beyond the school (Best, 1974; Kapenzi, 1972). Other approaches have been in the naturalistic tradition, focusing on the self-concept of the learner (Maslow, 1968). There have also been a variety of structured school programs in the formal tradition, ranging from DISTAR, Direct Instructional System for Teaching Arithmetic and Reading (Bereiter & Engelman, 1966) and the Peabody Kit (Dunh & Horton, 1966) at the pre-school level to programmed instruction (Skinner, 1954), IPI, Individually Prescribed Instruction (Glaser, 1968), and CAI, Computer Assisted Instruction (Atkinson, 1968) at the high school level. It is in this compensatory approach that the concept

of mastery learning was introduced by Bloom in 1968, with the claim that under appropriate instructional conditions virtually all students can learn to a high level most of what they are assigned to learn.

Supporters of mastery learning allege that unsuccessful and unrewarding school learning experiences can be countered within the context of ordinary group-based classroom instruction (Bloom, 1968; Block, 1971). Although there are a number of contemporary mastery learning strategies, according to Bloom (1974) there are four common characteristics: systematic instruction, feedback-correction, adequate time to achieve, and a clear mastery criterion. Furthermore, all begin with the assumption that such steps will help all, or almost all, reach a high level of achievement.

The idea of mastery is not new to American educational theory or practice, even though the specific terminology mastery learning was not coined until 1968 by Benjamin Bloom. Earlier examples of mastery learning with specific operational procedures are found in the Burk Plan of Individualized Instruction (Washburne, 1940), the Winnetka Plan (Washburne, 1922), the Morrison teach-test-reteach schema (Morrison, 1926), and programmed instruction (Skinner, 1954). However, it was Bloom who formulated the title mastery learning, a new and more dynamic slogan for an old idea.

The development of Bloom's mastery learning instructional model was related to the idea of compensatory education, of which Bloom (1965) was an early advocate. Compensatory

education became the watchword of the reform movement of the 1960s, in which the plight of the disadvantaged student, usually of a minority group status and poverty-stricken, was brought to national attention. Some of the most important federal legislation passed during this reform period included the Civil Rights Act and the Economic Opportunity Act, both of 1964, and the Elementary and Secondary Act of 1965. These federal measures had powerful implications for all levels of the public school, resulting in a myriad of compensatory programs for disadvantaged students (Gwynn & Chase, 1969).

One of the best known compensatory plans was the Head Start Program, created under the Economic Opportunity Act of 1964. The treatment of the disadvantaged child was first brought into focus by this pre-school summer program for poverty-stricken five and six year olds. A brief and intensive summer session was designed to provide children formally with what the middle class child received as a matter of course. However, most compensatory schemes, whether with Head Start at the elementary level or with Upward Bound at the secondary level, failed to meet their objectives. According to Findley (1972), compensatory programs such as Head Start exemplify an educational panacea, and it is his thesis that American education suffers most from oversimplified panaceas to inherently difficult problems.

The abandonment of "compensatory" terminology to avoid its condemnatory undertones was successfully achieved by Bloom in 1968 when he proposed mastery learning as a more

acceptable slogan. Nevertheless, the central focus of mastery learning did not change from the major objectives of compensatory education--how to help all pupils attain successful and rewarding school learning experiences. To some of those favoring excellence in education, mastery learning only represents another panacea in the history of American education which some educators erroneously believe "could lead education to its promised land" (Ebel, 1972).

Bloom derived the theoretical base of mastery learning from Carroll's (1963) conceptual model of school learning. In turn, most other contemporary mastery learning approaches derive their strategies from the Bloom model. Whether group-based or individually-paced, mastery strategies are built on the premise that the reason why many children, especially those from underprivileged backgrounds, do poorly in school is largely due to the lack of sufficient time and quality of instruction. Mastery procedures are designed to provide the learner with the needed time and a higher quality of instruction through feedback-correction procedures.

The assumption that all children exposed to mastery learning will benefit to the point of eventually becoming average, if not above average, in achievement ignores some significant aspects of the nature of differences in ability. The hope that the many children in the bottom half will equal many now in the top half is an unrealistic expectation (Biehler, 1971). In any mastery learning scheme, it appears very difficult to provide extra advantages for the bottom

group of students that will not also affect the top group. It seems that the same relative amount of opportunity which existed in the first place will be maintained, regardless of the mastery strategy. In fact, if mastery procedures provide a better quality of instruction, it is likely that the higher aptitude student will also benefit; thus the entire distribution of achievement will be shifted upward, leaving the same relative distribution.

Most mastery learning studies reported in the literature attribute increased performance to mastery strategies. However, there are earlier studies using a mastery format, without the label, which show contradictory findings. It is questionable, based on the large literature dealing with performance and individual differences (Cook, 1969; Conlon, 1970; Duchastel & Merrill, 1973), if any mastery strategy can ever homogenize performance. Even under optimum IPI (Glaser, 1968) and CAI (Atkinson, 1968) instruction, the most able students were able to achieve five times as much as the slowest students. These findings indicate that the less able students or the bottom half, achieve less than the higher aptitude students.

The present state of mastery research indicates that it has not systematically examined the various variables implicit in any learning arrangement. Variables of the original Carroll model as identified by Rice (1973) require systematic examination to establish evidence to confirm mastery learning allegations. Generally, proponents of mastery learning have

adopted a polemical approach rather than a research approach. The tone of the two books edited by Block (1972, 1974), a major proponent of mastery learning, is characterized by advocacy more than an experimental attitude. The studies are also single studies with no replication, so that the findings are based on a series of discrete studies rather than a systematic body of knowledge. Much of the mastery learning research as reported by Kim (1970, 1971) and Block (1972, 1974) is based upon crude comparisons of mastery with non-mastery procedures. While such studies point to the superiority of a mastery approach, the procedures are not sufficiently described in most cases to permit clear identification of the treatment variables.

One of the least considered variables in mastery has been aptitude, the attribute most central to the discussion of mastery learning and its relation to achievement by lower ability students. If mastery learning is to function as a corrective or compensatory mechanism for lower aptitude students, it is essential that the mastery procedures utilized not merely raise mean performance of a class, but specifically raise the performance of lower aptitude students. Furthermore, it is necessary to know what criterion level facilitates achievement of lower aptitude students. A high level of mastery, e.g., 90%, might not be as effective as a lower level, e.g., 70%, if it discourages pupil participation. Presently there is no evidence to indicate the optimal procedural elements of presentation and correction by subject,

age, or aptitude. Consequently, it was decided to examine the aptitude of students from the standpoint of differential criterion levels and the interaction effect of these two independent variables upon student achievement, retention, and attitude. The objectives and research hypothesis of this study are specified below.

Objectives and Research

Hypotheses

This study had several objectives. The first major objective was to ascertain if different criterion levels helped children perform at a higher level as measured by the number of responses on an achievement test. The second major objective was to ascertain the relation of vocabulary level to treatment criterion. These major purposes contributed to the selected objectives as stated below:

1. To test the effect of 90, 80, and 70% criterion levels on:
 - a. student achievement,
 - b. retention, and
 - c. attitude.
2. To test the effect of vocabulary on:
 - a. student achievement,
 - b. retention, and
 - c. attitude.
3. To test the interaction of vocabulary and treatment on:
 - a. student achievement,
 - b. retention, and
 - c. attitude.

There were two independent and three dependent variables in this study. The two independent variables were criterion level and aptitude. For purposes of this study, these two terms are defined as follows: criterion level is a predetermined percentage of correct responses required for mastery; aptitude refers to an achievement score on a vocabulary test (See "Definition of Terms," p. 18).

The term "mastery" would seem to imply a perfect performance level of 100% or a high performance level of 95%. Evidence indicates that even a select college population represents the requirements of perfect mastery in a learning sequence (Sherman, 1967; Healey & Stephenson, 1975). Block (1970), however, is the only investigator who has examined the affect of different criterion levels--95, 85, 75, and 65%--on various cognitive and affective variables. He found that the highest level of 95% facilitated achievement but that the 85% level contributed to a better mix of sustained interest in the subject and achievement (See "Review of Literature," p. 65). Block was not concerned with the problem of the low aptitude student, a major focus of this study, and did not use an aptitude measure as a concomitant variable.

The investigator therefore selected the 90, 80, and 70% levels for this study on the basis of experience and reference to other studies. It was known beforehand that the sample in this study would consist of a large number of low performing students, as measured by a vocabulary score. Experience as a teacher with low aptitude students suggested

that the investigator not select a criterion level so high that it would adversely affect motivation and depress interest in the study. A high criterion level of 90%, slightly below the highest used by Block, seemed to be a good intermediate level in terms of his findings as well as conforming to the common criterion used in below-college level studies of mastery learning. Even though at first glance higher levels might appear to be appropriate, the investigator concluded that the 90, 80, and 70% levels were reasonable levels for the objectives of this study. Criterion level, the treatment variable of the study, was utilized to determine the extent, if any, to which differential performance levels facilitated achievement, in particular for low aptitude students.

The second independent variable, aptitude, was considered a critical variable of this study. Mastery learning proponents claim that the maintenance of a high criterion level on formative exercises help lower ability students to overcome learning difficulties, customarily attributed to low aptitude, and to attain high levels of achievement on summative posttests. While studies comparing mastery with non-mastery procedures report an increased number of students achieving to higher performance levels, there is a lack of evidence to substantiate the claim of mastery learning that mastery procedures overcome achievement deficiencies related to low aptitude (See "Review of the Literature," pp. 53-58).

The three dependent variables were achievement, retention, and attitude. Achievement, the first cognitive criterion, is most frequently used by schools to measure learning (Ebel, 1972). Data on achievement reported by Collins (1970), Kim (1970), Block (1971), and Kersh (1971) point to an increased percentage of students attaining high summative achievement levels when required to maintain selected performance levels throughout highly sequential mathematical content. While these studies suggest a strong relationship between the maintenance of a prescribed criterion level during sequential learning and final achievement, only Block (1971) has investigated which performance levels maximize achievement for eighth graders studying matrix algebra. This researcher thought it important to ascertain in certain performance levels facilitated learning for lower aptitude students studying social studies, a less structured content area than mathematics.

The second cognitive criterion, retention, is defined by Brownell (1948) as the maintenance of skills or knowledge with no practice after the learning's completion. Block (1970), Kersh (1971), and Romberg, Shepler, and King (1970) reported a significant relationship between the attainment of certain performance levels and retention. These studies suggest that requiring learners to achieve to performance levels they might otherwise not attain will have a positive affect on retention. Also, Block's findings indicate that particular levels may maximize retention. It was decided to examine

the relationship between the maintenance of certain criterion levels and retention, as well as the interaction effect of those two variables on students of varying aptitudes, especially those of lower ability..

It is conceded in the literature that there is a relation between treatment and attitude (Sherman, 1967; Sheppard & McDermot, 1970). It suggests that the attainment of a high criterion level is related to a feeling of the student that he is achieving more adequately (Feather, 1966). For example, if a learner is required and assisted to maintain a criterion level throughout his learning that he might not attain given his previous learning predispositions, a positive change toward learning might result from the fact that he learned much better than expected. However, the effect of stringent performance levels on intermediate grade students of lower ability is still unknown. Therefore, an investigation of attitude was undertaken to determine the relation of differential criterion levels to attitude toward the treatment.

Unit completion time was a fourth dependent variable originally considered by the investigator, as indicated in the request to the El Paso Public Schools, Appendix G. Only a few of the children in the pilot phase of the study failed to complete the unit within the allotted time, giving the investigator the impression that the 15-day time period was adequate for the mastery procedures, described in Chapter III. However, during the experimental study, more children than had been predicted failed to complete the unit, especially

in the low aptitude group where the mean lesson completion rate for almost one-third of the students was only 23 of the 44 unit lessons. Although self-report data on unit completion time had been collected, the fact that many students failed to finish all the lessons made such data meaningless. Unit completion time was therefore dropped as a dependent variable in the study. Mastery procedures require that time be open ended, so that lower aptitude students have an opportunity to utilize as much time as needed to achieve the criterion level set for mastery. Therefore, the investigator notes that a fixed instructional period automatically precludes the operation of mastery procedures for all students.

From the objectives of the study, the following research hypotheses follow:

1. There is a significant difference ($p < .05$) by criterion level--90%, 80%, and 70%--on
 - a. achievement, as measured by a geography posttest,
 - b. retention, as measured by a delayed posttest, and
 - c. attitude toward the unit, as measured by the Attitude Toward Any Subject scale (Remmers, Short A Form).
2. There is a significant difference ($p < .05$) by aptitude level, as measured by the vocabulary sub-test of the Iowa Tests of Basic Skills, Form 6, Level 13, on
 - a. achievement,
 - b. retention, and
 - c. attitude.
3. There is a significant difference ($p < .05$) in the interaction of treatment criterion level and vocabulary on

- a. achievement,
- b. retention, and
- c. attitude.

The next section defines terms as used in this study. Most of the terms follow stipulative definitions; operational characteristics are enlarged upon in Chapter III Procedures.

Definition of Terms

The following definitions indicate how the terms were used in this study.

Mastery Learning. The operational characteristics of mastery treatments vary since each investigator stipulates his own procedures. However, mastery learning in general describes a teach-test-reteach strategy with a feedback-corrective component to help students overcome learning problems and achieve mastery. In this study mastery procedures included diagnosis, correction, and restudy of content if criterion were not attained. Upon completion of the review procedure, all students proceeded to the next unit lesson whether or not the criterion level had been reached. The specific operational procedures are discussed in Chapter III.

Criterion Level. The criterion level is a predetermined performance standard which represents a proportion of right responses on each lesson and review test. Three criterion levels were used, the 90%, 80%, and 70% levels.

Aptitude. Aptitude represents a score on the word meaning section of the Iowa Tests of Basic Skills: Form 6, Level

13 (Lindquist & Hieronymus, 1971) used to determine a high, middle, or low vocabulary group for purposes of blocking on vocabulary as a concomitant variable.

Vocabulary Level. Vocabulary level refers to a high, middle, or low group to which students were assigned for purposes of statistical analysis on the basis of the word meaning section of the Iowa Tests of Basic Skills: Form 6, Level 13 (Lindquist & Hieronymus, 1971).

Achievement. Achievement is the score made on the unit test for Population Growth in Mexico and the United States administered as a posttest upon completion of instruction.

Retention. Retention is the score made on the same form of the unit test administered as a delayed posttest three weeks after instruction.

Attitude. Attitude is a negative or positive disposition toward the treatment as measured by an Attitude Toward Any Subject scale, Remmers, Short A Form (Silance & Remmers, 1934).

Chapter II, the Review of the Literature, discusses the antecedents of mastery learning. Two major mastery approaches, from which most contemporary strategies are derived, are explained in the second part of the review. The third part examines mastery learning research. The last section describes the origins of the present study.

CHAPTER II

REVIEW OF THE LITERATURE

This chapter is divided into four parts. First, the antecedents of mastery learning are discussed briefly since a mastery approach to education is not a new idea in American education. The second part explains two major mastery learning approaches, the Bloom, or class paced, strategy and the Keller, or individual paced, strategy. The bulk of reported mastery research has employed one of these two approaches. The third part reviews previous mastery research. The final part describes how this study originated from an earlier mastery study by Block (1970) who examined differential criterion levels.

The History of Mastery Learning

Recent precursors of mastery learning include Burk's plan of individualized instruction, Washburne's Winnetka Plan, Morrison's teach-test-reteach paradigm, programmed instruction, and Carroll's Model of School Learning. It was not until 1968, however, that Bloom formally introduced the terminology "mastery learning."

Frederic Burk of the San Francisco State Normal School introduced the Burk Plan of individualized instruction in the early 1900s. His teaching plan was an attack against the traditional style which he compared to the military system and termed "Lock-step Schooling" (Washburne, 1940). As to the effectiveness of the Burk Plan, Washburne stated (1940, p. 251):

Schools that have tried Burk's plan intelligently have found that it worked, that it saved time, that it reduced retardation, that it made school life happier and more efficient for teacher and pupils, that it actually made possible the wide range of individual difference in readiness among children in a given class.

In the early 1920s Carleton Washburne and his associates introduced the Winnetka Plan, one of the first comprehensive efforts made to individualize classroom instruction (Kersh, 1965). Washburne credited Burk for providing the basis of the Winnetka Plan. Under this plan units of achievement replaced traditional time units with promotions being based on individual achievement rather than time. Failure, grade repetition, and grade skipping were eliminated as a child proceeded to master units of achievement, called goals, at his own pace. Each goal had to be mastered by each pupil before proceeding to the next (Washburne, 1922).

The features of the Winnetka Plan of individual work consisted of three basic steps. The first step called for

the establishment of definite goals of subject matter units. Second, there was the preparation of tests which completely covered each subject matter unit and diagnosed the difficulties of individual children. Third, self-corrective practice materials enabled children to both prepare for these tests and make up the deficiencies shown by the tests.

During the 1920s Henry C. Morrison (1926) developed another precursor of mastery learning at the University of Chicago's Laboratory School. Morrison's "mastery formula" consisted of the following steps under a group-based, rather than individualized, method of instruction: pre-test, teach, test the result, adapt procedure, and teach and test again to the point of mastery.

The pre-test procedure served two important purposes. First, it oriented the teacher and gave him ground for an intelligent approach to the particular problem before him; and, second, it tended to establish in the pupil's mind a connection between prospective learning and present attainment. In a few cases pupils were excused from studying the unit for already having mastered the content material.

The testing aspect of the procedure also had two purposes. First, it provided information as to whether or not the child should proceed to the next unit. Second, it helped the teacher decide in which manner to modify the teaching procedure if the test disclosed that mastery had not been achieved. Reteaching took place as many times as a teacher judged necessary to bring most, if not all, students to

mastery level. Pupils unable to respond to the routine instruction of the class group as a whole received special study and treatment.

The Burk, Winnetka, and Morrison plans shared five major features. One, mastery was defined in terms of specific educational objectives which each pupil was expected to attain. Two, course work was organized into a sequence of learning units. Each unit contained specific unit objectives and a collection of learning material to teach those objectives. Three, students were required to master each unit before proceeding to the next. Four, a feedback-corrective component provided teachers and students with diagnostic information regarding the adequacy of a student's learning. Diagnostic tests indicated either unit mastery or material still to be mastered. Students failing to achieve unit mastery were provided with supplementary material. Finally, each plan treated time as a dependent variable. Under the Burk and Winnetka plans, student learning was individually-paced; each pupil was allowed all the time he needed to master a unit. The Morrison method allowed each student the learning time his teacher required to bring all, or almost all, students to unit mastery under group-paced instruction.

Washburne (1940) gave the following reasons as to why such mastery instructional plans failed to spread more widely by 1940:

- 1) a lack of adequate tests and texts for individual work,

- 2) a shift in emphasis from subject matter to child-centered learning, and
- 3) the development of "compromise plans" of ability grouping and group projects in which each child participated according to his own ability.

Programmed instruction was the next major antecedent of mastery learning procedures. Pressy (1926) is usually considered the first person to make a systematic effort to automate classroom instruction. What Pressy heralded in 1930 as the "coming industrial revolution in education" failed to materialize, however, and it was not until the 1950s that Skinner's application of reinforcement theory to programmed instruction brought about a revival of interest in programmed learning (Keller, 1968). The basis for programmed instruction is also a component in mastery learning. Programmed learning begins with the idea that the learning of any complex behavior depends upon the learning of a sequence of less-complex component behaviors (Skinner, 1954). Theoretically, a student can master any complex behavior so long as it is broken down into a chain of component behaviors. If a student can master each link, it is then possible for him to learn even the most complex skills.

Operationally, programmed instruction involves the sequential presentation of material and a response requirement for each stage. Subject matter is broken down into steps called "frames" which require some overt response. Immediate

feedback to the student indicates mastery or non-mastery of the behavior or frame presented. An incorrect response is immediately corrected before misunderstanding results. A correct response brings a reinforcing positive confirmation before proceeding to the next frame.

Increasing interest in programmed instruction led to large-scale field-testing of programmed instruction in various types of schools throughout the country. Among the first school systems to investigate programmed instruction in large-scale terms were the Roanoke city schools of Virginia (Rushton, 1963), the Denver Public Schools of Colorado (Jones, 1962; Reed, 1962), and the Rural schools of Utah as part of the Western States Small School Project (Ford & Walker, 1961).

The initial enthusiasm for programmed instruction in the 1950s was followed by disappointment in many circles as some of the limitations of a primarily technological approach to education became evident. First of all, many publishers hurriedly produced ineffective written materials in an attempt to beat their competitors. Secondly, the cost of more elaborate machines was so great that school systems were unable to buy them without some type of subsidy. Most government and foundation grants were reserved for pilot studies (Biehler, 1971).

Despite the noted limitations associated with programmed instruction, it has been used successfully with a variety of students from pre-kindergarten upward and with the retarded, the average, and the gifted (Schramm, 1964; Stolurow, 1962).

Block (1971) noted in a review of mastery learning history that programmed instruction provided an especially valuable tool for those students who required small learning steps, drill, and frequent reinforcement; however, programmed instruction only provided a useful mastery learning model in situations not constrained by fixed classroom groupings. In effect, programmed instruction operationally provides the basis for an individual based approach to mastery learning, as demonstrated in large scale studies (Atkinson, 1968; Glaser, 1968) and in social science subjects (Thomas, 1967; Fishburne, 1971). Overall, programmed instruction has had limited acceptance in comparison with regular texts and non-print materials. There has also been limited acceptance of programmed instruction by school administrators and teachers.

It was John B. Carroll's "Model of School Learning" of 1963 that served as the basis for the more useful mastery models to emerge in the late 1960s (Bloom, 1968; Keller, 1968). Even though Carroll did not identify himself with the mastery movement until 1971, his proposed "Model of School Learning" in 1963 consisted of a conceptual paradigm outlining the major factors influencing student success in school learning. Since this model has become to be theoretically identified with the mastery learning movement, it is important to describe the nature of the model.

According to Carroll, the degree of learning is a function of five interrelated factors. The full Carroll model is:

Degree of Learning = f $\frac{\text{Time Spent}}{\text{Time Needed}}$ (1. Time Allowed 2. Perseverance)
 (3. Aptitude 4. Quality of Instruction 5. Ability to Understand Instruction)

The Carroll Model of School Learning places an emphasis upon time as dimensions of school learning. First, there is the amount of time a student actually spends in instruction. This is usually a combination of the two factors--time allowed and the amount of time the student is actively involved in learning. Second, there is the time actually needed. Time needed is affected by the aptitude a student brings to the learning task, the quality of the instruction, and the student's ability to understand instruction. Aptitude is an index of not only the level to which a pupil will learn in a given time, but also the amount of time needed to learn to a given level under optimal learning conditions. Quality of instruction is defined in terms of the degree to which the presentation, explanation, and ordering of the learning task's elements approaches the optimum for each individual learner. Ability to understand instruction represents the student's ability to generally profit from the instruction and is closely related to general intelligence. Carroll proposes that the quality of the student's instruction and his ability to understand it interact to extend the time needed for task mastery beyond that normally required by aptitude for the task. If quality of instruction and ability to understand it is high, then little or no additional learning time is required to master a learning task. However, if they are both

low, then much more additional time is required for task completion. The Carroll model is especially pertinent to conceptualizations of mastery learning because it theoretically emphasizes the need for additional time for lower aptitude students. But when a mastery procedure, whether for deficiency of operational procedures or defect in student perseverance, fails to involve the student in more time in learning, the mastery procedure cannot work as a remedial and corrective strategy because the necessary component of additional student time input is absent.

A major criticism of the Carroll model is that it is presented as an equation when, in fact, there is no complete network of equations connecting the various components of the model. Major measurement problems are associated with the model of school learning; for example, perseverance is perhaps the most difficult component of all to measure or to predict. In fact, Carroll shifted his emphasis from the measurement of perseverance to its enhancement through teacher praise and reward. Equally difficult to measure is the quality of instruction. With respect to teacher ability to manage instruction, which according to Carroll is synonymous with quality of instruction, there has been no systematic investigation of the relationship of teacher characteristics to pupil mastery. There has also been no systematic study of material variables. Another difficult measurement task is related to the true amount of time that a pupil needs to learn because this is a variable that cannot be observed

directly since it assumes that a student is well motivated to spend all of the necessary time needed to learn and that instruction is optimal. Despite the major measurement problems associated with Carroll's model of school learning and its tautological premise, this model has been the basis of contemporary mastery strategies.

It was in 1968 that the construct of mastery learning was formally introduced by Benjamin Bloom in a seminal article "Learning for Mastery," adapted for the 1971 Block edition as "Mastery Learning." Bloom explicitly placed his rationale for mastery learning on the five factors of Carroll's "Model of School Learning," previously discussed. However, Bloom also acknowledges the eclectic sources of the idea in the work of a number of writers, including Morrison and Skinner, previously discussed. Bloom goes much farther than previous authors and suggests that remedial strategies, hereafter referred to as mastery learning, are potentially so powerful that "ninety-five per cent of the students . . . can learn a subject to a high level of mastery (for example, an A grade) if given sufficient learning time and appropriate types of help" (In Block, 1971, p. 51). In this respect, Bloom depreciates the importance of aptitude, as traditionally interpreted, and places himself in the school of the "environmentalists," citing his own work (Bloom, 1964) and that of Hunt (1961) as optimistic grounds for believing that manipulation of the conditions of learning are more important than aptitude. He twists the studies of Atkinson (1968) and

Glaser (1968) dealing with individual programmed instruction as evidence that most students can attain a given criterion of achievement. What he deemphasizes is the fact that, in the time given, highest aptitude students learned five times faster than the lowest aptitude students. Nevertheless, he recognizes that a basic problem of mastery learning is to find ways to reduce the learning time of lower aptitude students so that the task will not be excessively long. He also makes an admonition that is worth quoting: "It is unlikely that mastery can be attained in a given term by students who have had a long history of learning difficulties in such subject" (In Block, p. 55). Implicit in this statement, taken in the context of Bloom's general approach to compensatory education, is the idea that mastery procedures must begin early in the life of the child and be sustained over a long period of learning. Short-term intervention, the kind usually involved in mastery learning experiments, might bring about experimentally significant results which might be regarded as purely fortuitous in the context of general learning.

Bloom explicitly recognized that there were a number of mastery alternatives; however, he described in some detail the procedures of mastery learning supplementing regular group instruction. Consequently, the terminology "Bloom mastery strategy" has become identified with group paced mastery; an individual paced strategy has been particularly identified with Keller, and have become known as "Keller strategies." These two mastery strategies are described in the next

section. This review of the historic antecedents of mastery learning indicate that there were mastery models used before the term was introduced. In fact, the whole concept of mastery is concisely summed up in the model of Morrison, which is to "test, reteach, and test again."

Bloom and Keller Mastery

Learning Strategies

Two major models of mastery learning are Bloom's "Learning for Mastery" model and Keller's "Personalized System of Instruction." One or the other of these serve as a basis for most contemporary mastery strategies.

Bloom's mastery learning model is predicated upon the assumption that up to 95% of students can learn much of what they are taught to the same high levels customarily reached by the best students. The problem is to be able to define what is meant by mastery of a subject and then to be able to provide each student with the time and quality of instruction needed to demonstrate this mastery. Bloom's mastery strategy is primarily designed for use in a group-based instructional situation where the time allowed for learning is relatively fixed, although the basic ideas are equally applicable in an individual-based instructional situation.

The first step in Bloom's mastery strategy entails the formulation of an entire set of cognitive objectives that all students will be required to achieve to a prescribed level by the course's end. A course is then broken down into a

sequence of smaller learning units, each typically covering the objectives contained in about two weeks' instruction.

The material in one unit builds directly upon the material in the previous units. A sequence of learning tasks are constructed for each unit to serve as a blueprint on how each unit might be taught for mastery.

Once the learning tasks for each unit have been described, a brief, ungraded, and student-scored diagnostic-progress test called a formative evaluation is constructed for each segment or unit (Bloom, Hastings, & Madaus, 1971). The formative evaluation, given at the end of each unit, indicates which of the objectives within the unit a student has failed to grasp. Finally, the teacher prepares a set of alternative learning materials or instructional correctives keyed to the formative tests. These correctives teach the unit objectives, but they do so in ways different from the teacher's group-based presentation. Learning corrective procedures include small group study sessions, tutorial assistance, review of the original instructional materials or alternative textbooks, workbooks, programmed materials, audio-visual materials, or other materials.

The Bloom strategy calls for the teaching of the first unit of the mastery sequence under a normal group-based instructional situation. Upon completion of the first unit, a formative evaluation is administered as a checking device to see how each student is achieving unit goals. Those who do not achieve as desired are referred to instructional

correctives in order to attain the unmastered objectives. A student may complete his learning on his own time or else be provided with class time for review before group instruction resumes on the next unit of instruction.

The cycle of group-based instruction, formative testing, and prescription-corrective procedures for each learner on each unit is followed until all instructional units have been completed. The course final examination, called a summative evaluation, is administered at the conclusion of instruction. The summative test measures achievement with respect to overall course objectives.

Bloom defines mastery operationally as performance at or above a particular level (usually 80% to 90% correct) on the final examination. The student's score on the summative evaluation determines the final grade. Regardless of the number, any student who performs at or above the stipulated mastery level earns an A grade. Lower grades are given to those performing below mastery. However, Bloom alleges that only a handful of students should obtain a B or C if the Bloom-strategy has been followed by the instructional leader.

Keller's (1968) "Personalized System of Instruction" approach to mastery learning is a second major mastery model. The Personalized System of Instruction (PSI) is essentially programmed instruction whereby the frames are substantially enlarged and a personal-social aspect is introduced.

An instructor begins the PSI procedure by predefining the cognitive course objectives. These objectives are then

subdivided into a number of learning units requiring about one week of work to master. Procedures whereby a student masters each unit typically include a list of the objectives, a suggested set of study procedures based heavily on textbooks or other materials, a set of study questions, and a set of test items over the unit's objectives.

A student proceeds through the units at his own pace. After each unit is completed, a student is administered a unit examination by a proctor or a teacher. The examination, upon completion, is corrected immediately by a proctor or teacher. Perfect performance means that a student proceeds to the next unit. The test, review, retest cycle is continued until a student is able to demonstrate perfect performance on a unit.

Mastery is operationally defined by Keller as perfect performance on a particular number of units by a certain point in time. Grades are usually determined by the number of units completed by a particular point in time.

The two major approaches described above commonly seek to help all, or almost all, pupils attain a high level in their learning. Both strategies require prespecification of cognitive objectives which each student will be expected to achieve to some high level. Both require that a course be broken into a sequence of smaller learning units. The first part of a unit is an instructional component, and the second part consists of a feedback-correction feature. The purpose of this latter feature is to monitor the effectiveness of

the original instruction of each student and provide appropriate corrective actions. Finally, grades are determined solely on the basis of a pupil's absolute performance rather than on his relative performance. This grading procedure does not differ from traditional criterion-referenced testing procedures for classroom instruction. It is the feedback procedure of the mastery method which allegedly produces higher achievement results than traditional classroom teaching methods.

Keller procedures, which rely on more pupil initiative and responsibility, are assumed to be more appropriate for the college level while Bloom procedures, which involve more teacher direction, are assumed to be more suitable for younger learners. Because Keller mastery grades are cumulative unit rather than end-of-course grades, there is a built in tendency for Keller mastery grades to exceed non-mastery comparisons. The next section will briefly summarize mastery research pertinent to elementary instruction.

Mastery Learning Research

This review of mastery learning research is classified on the basis of five variables--achievement, retention, attitude, time, and aptitude. Mastery research related to the social studies is also discussed. The purpose of the classification is to demonstrate the trends of previous mastery research and the almost complete absence of studies that examine the various mastery variables implicit in any

learning system, especially the relationship of aptitude to achievement under differential criterion levels.

In reviewing the literature, the investigator has identified 51 studies of mastery learning. Of these studies, 35 were at the college level, and generally were rough comparisons of mastery versus non-mastery treatments (these studies are listed alphabetically in Appendix G). It is inappropriate to generalize from mastery treatments at college level to mastery treatment at grade school level because of age and aptitude selection factors. This review of the mastery literature, therefore, will be restricted to mastery learning studies below college level, except where variables of particular interest to this study, in addition to achievement, have been investigated, i.e., retention, attitude, time spent in mastery, and aptitude.

Achievement

Fourteen studies reporting achievement for mastery and non-mastery strategies are summarized in Table 2.1. Thirteen of the studies were at the upper elementary or junior high levels; only one was at a primary level. Eleven of the studies were in one subject; three--two by Kim and one by Lee--were in more than one subject. Nine studies were in arithmetic or mathematics; two in English, three in science, one in foreign language, and four in social science. The four social science mastery studies (Fagan, 1975; Gaines, 1971; Jones, 1974; and Wyckoff, 1974) were conducted under

Table 2.1
Achievement Under Conditions of Mastery and Non-Mastery Learning

Study and Year	Number in Sample	Level and Content Area	Duration	Results
Block, J. H., 1970	91	8th, Matrix Algebra	5 days	85% & 95% criterion groups achieved significantly higher than non-mastery students on summative posttest.
Collins, K. M., 1970	6 treatments 25 per class	8th, Mathematics	N.S.	Objectives, diagnostic problems and review prescriptions increased the number attaining criterion from 40% to 80%.
Fagan, J., 1975	470	7th, Geography	4 weeks	No significant difference on posttest between mastery and non-mastery groups.
Gaines, G., 1971	28 classes	5th-8th, Anthropology	5 weeks	No significant difference between mastery and non-mastery.

Table 2.1 (Continued)

Study and Year	Number in Sample	Level and Content Area	Duration	Results
Jones, G., 1974	20 classes	7th, Geography	4 weeks	Only middle aptitude groups of mastery treatment achieved significantly higher than non-mastery group.
Kersh, M. E., 1971	5 classes	5th, Arithmetic	1 year	63% of exp. group reached criterion compared to 14% in the control group on the final exam.
Kim, H., 1969	273	7th, Arithmetic	N.S.	74% of exp. group attained criterion on final test compared to 40% of control group
Kim, H., 1970	5,800	7th, English, Mathematics	8 weeks	Percent Achieving Mastery on Summative Test <div> <div>Subject</div> <div>English</div> <div>Math</div> </div> <div> <div>Exp.</div> <div>72%</div> <div>61%</div> </div> <div> <div>Control</div> <div>28%</div> <div>39%</div> </div>

Table 2.1 (Continued)

Study. and Year	Number in Sample	Level and Content Area	Duration	Results
Kim, H., 1971	25,887	7th, English, Mathematics, Physics, Biology	N.S.	Percent Achieving Mastery on Summative Test English Math Physics Biology Mastery 50% 48% 30% 22% Non-Mas- 44% 26% 85% 37% tery
Lee, et al. 1971	12,504	5th & 6th Mathematics, Science	N.S.	Percent Achieving Mastery on Summative Test Subject Grade Exp. Control Math 5 42% 12% Math 6 45% 18% Science 5 39% 7% Science 6 46% 12%
Romberg, A., Shepler, J., King, I., 1970	N.S.	6th, Mathe- matical Proof; and Statis- tics	N.S.	Mastery group achieved significantly higher than non-mastery group.

Table 2.1 (Continued)

Study and year	Number in Sample	Level and Content Area	Duration	Results												
Smith, M. I., 1968	75	6th Spanish	N.S.	A comparison of listening comprehension scores of one mastery group with two control groups revealed significantly different mean scores. Mastery: 76.8 Control A: 62.6 Control B: 62.3												
Walbesser, H., and Carter, H. 1968	N.S.	1-3, Science	N.S.	Percent Achieving Mastery on Summative Test <table><tr><td>Grade</td><td>Mastery</td><td>Non-Mastery</td></tr><tr><td>1</td><td>88%</td><td>54%</td></tr><tr><td>2</td><td>58%</td><td>19%</td></tr><tr><td>3</td><td>64%</td><td>22%</td></tr></table>	Grade	Mastery	Non-Mastery	1	88%	54%	2	58%	19%	3	64%	22%
Grade	Mastery	Non-Mastery														
1	88%	54%														
2	58%	19%														
3	64%	22%														
Wyckoff, D. B., 1974	4 classes	6th, Anthropology	10 weeks	No significant difference between Mastery and non-mastery treatment.												

the auspices of the Geography and Anthropology Curriculum Projects, University of Georgia, although Wyckoff was a student at Georgia State University. The duration of studies ranged from five days (Block, 1971) to eight weeks (Kim, 1970). The length of seven studies was not reported.

Except for the four mastery studies in social science, the reports of findings seem to indicate that mastery procedures facilitated achievement in comparison with control or non-mastery procedures. The data is not reported in the same manner for all studies, and thus it is difficult to draw firm conclusions. As Table 2.1 indicates, most of the studies report the percentage of mastery and non-mastery students attaining the criterion of mastery. However, the four social science studies report data in terms of significant difference between mastery and non-mastery groups. Although details are lacking, it might be inferred that hierarchically sequenced subjects such as arithmetic lend themselves more readily to mastery procedures than do social science subjects in which the sequencing of material may be logical but not necessarily hierarchical. This aspect of mastery learning will be dealt with in more detail in the section critiquing the social science mastery learning studies. In general, the reports of mastery and non-mastery in Table 2.1 are crude comparisons of mastery and non-mastery procedures. The reports of the treatment and experimental procedures in many cases are too scanty to make firm conclusions about the efficiency of mastery treatment,

notwithstanding the tendency to show that feedback-correction procedures facilitate the attainment of mastery.

Retention

Retention has been explored in five mastery studies below college level, one each at grades five, six, and eight and two at seven. Three were in arithmetic or algebra; two were in geography. The delay between the achievement test and the delayed posttest was usually rather short, about 14 days, as in the Block study (1970) but was about three months in the Kersh study (1970).

Retention is an important concept in educational experimentation. The object of instruction is not merely that the student be able to perform a learning task immediately after instruction, but that he retain the knowledge or skill to perform the learning task at some future time. The main problem of learning academic subjects, according to Ausubel (1968), is the problem of maintaining the availability of previous learned knowledge. The problem of forgetting characterizes all learning, especially verbal factual material which is not used between the time of initial learning and the time the student is called upon to reproduce it, usually on a test. Consequently, experimental treatments which are designed to improve instruction frequently involve a retention test. The superiority of organization, lucid presentation, correction, review, and pacing manifest itself at some delayed time after instruction; so runs the theory,

irrespective of the construct of mastery (Ausubel, 1968).

In the five studies in Table 2.2, the superiority of mastery treatment was manifested in the three arithmetic-mathematics studies (Block, 1970; Kersh, 1971; and Romberg, et.al., 1970) but not in the two geography studies (Jones, 1974; Fagan, 1975). In the achievement testing in these studies there was a similar finding (See Table 2.1). The five studies of mastery using a retention measure would not permit one to conclude that mastery strategies always facilitated retention; more studies in different subjects at different grade levels would be required. The three mastery studies in arithmetic and algebra which show a retention advantage for the mastery treatment are simply consistent with the body of general research which indicates that there appears to be less forgetting in hierarchically organized arithmetic-mathematics than in verbal factual material.

Function of Cities (Jones, 1974) and Transportation (Fagan, 1975) were the subjects of the two geography mastery studies. Both of these texts require the student to learn and retain a large amount of discrete factual material. Since social studies is characterized by a great deal of factual material which does not carry over from one period, area, or topic to another, learning in social studies is particularly subject to forgetting. Probably, the only way for mastery in social science to demonstrate a retention advantage would be to demonstrate a quantitative learning

Table 2.2

Retention Under Conditions of Mastery and Non-Mastery Learning

Study and Year	Number in Sample	Level and Content Area	Duration	Post Test Delay	Results
Block, J. H., 1970	91	8th, Algebra	5 days	14 days	85% & 95% criterion groups retained significantly higher than 75% or 65% on the non-mastery groups.
Fagan, J., 1975	470	7th, Geography	4 weeks	3 weeks	No significant difference between mastery and non-mastery.
Jones, F. E., 1974	20 classes	7th, Geography	4 weeks	17 days	No significant difference between high and low aptitude subjects of experimental group and those of control group. Middle aptitude group of mastery treatment retained significantly more than non-mastery.
Kersh, M. E., 1970	5 classes	5th, Arithmetic	1 year	summer vacation	63% of mastery group earned 80% correct responses on delayed posttest.

Table 2.2 (Continued)

Study and Year	Number in Sample	Level and Content Area	Duration	Post Test Delay	Results
Romberg, A., Shepler, J., and King, I., 1970	N.S.	6th, Mathematical Proof; Probability and Statistics	2 weeks 4 weeks	14 days 28 days	Correlation between achievement and retention $+ .75$ and $+ .78$ on each unit, respectively. Mastery students retained significantly more than non-mastery.

superiority on the achievement test. Thus while mastery students would forget, they would retain more simply because they had learned more in the first place. In the absence of any demonstrated mastery superiority on the achievement test, one should not anticipate a mastery superiority on a retention test in a social studies subject.

Attitude

Only five mastery learning studies have investigated the relationship of mastery to student attitude, as shown in Table 2.3. Of these studies, only that of Block (1970) was conducted at the elementary level and explored different criterion levels in relation to affective measures. He found that while the 95% criterion level on formative tests increased summative performance, the 85% level produced sustained high interest in and attitude toward the subject over time. He therefore concluded that a mid-criterion level, neither too high nor too low, best met the dual objectives of raising performance and maintaining a positive attitude toward instruction. A difficulty with the Block finding, however, is the short duration of the study, one week, including data collection time.

The four other studies reporting attitudinal outcomes provide contradictory findings. Sheppard and MacDermont (1970) reported higher student interest in a mastery than in a non-mastery course in college psychology but Lawler, Dick, and Riser (1974) found no difference in attitude toward a

Table 2.3

Attitudinal Outcomes Under Conditions of Mastery and Non-Mastery Learning

Study and Year	Number in Sample	Level and Content Area	Duration	Results
Block, J. H., 1970	91	8th, Algebra	5 days	85% & 95% levels had more positive effects, but 85% level produced sustained high levels of interest and attitude toward subject over time. Achievement scores 85% group increased, but attitude scores dropped.
Healy, J. R., and Stephenson, L. K. 1975	N.S.	College, Introductory Geography	semester	Informal observational data; no quantifiable. Initial novelty effect gave way to dissatisfaction attributed to early need for systematic study. By fourth week complaining diminished, perhaps because students made adjustment to regular study.
Lawler, R. M., Dick, W., and Riser, M., 1967	167	College Health Education	quarter	No treatment mean difference on attitude questionnaire toward health education. Scores almost the same for the traditional type instruction and the computer managed instruction.

Table 2.3 (Continued)

Study and Year	Number in Sample	Level and Content Area	Duration	Results
Sherman J., 1967	N.S.	College Introductory Psychology	semester	Final reaction to course was positive as measured by questionnaire, but initial resentment toward perfect per unit mastery requirement was noted.
Sheppard, W., and Mac Dermot, H., 1970	300	College Introductory Psychology	N.S.	Levels of student interest in attitude toward the topic learned significantly higher for mastery group.

college health education course by mastery and non-mastery treatment groups. Sherman (1967) and Healey and Stephen (1975) found negative attitudes toward perfect mastery required in a college psychology course giving way to more positive attitudes toward the end of the course.

The mastery learning research associated with affective variables indicate not only a paucity of research but inconclusive findings. Since a favorable attitude toward instruction may be related to perseverance in the mastery learning task, it was decided to collect attitudinal as well as cognitive data in the present study.

Time

One of the most significant variables in the Carroll Model of School Learning, previously discussed, is that of time. In fact, a short form of the Carroll Model expresses learning as a function of time spent in learning in relation to time needed to learn (See page 27). One of the dilemmas in school instruction is that the lower aptitude students, who actually need to spend more time in learning a task than do higher aptitude students, frequently spend less time in the learning task than do higher aptitude students. Differences in learning attributed to aptitude become even greater because of differences in time spent in the learning task. Consequently, part of the strategy of mastery learning becomes the devising of feedback and corrective

procedures which require that the time spent in learning more nearly approximates the time needed to learn.

Notwithstanding the importance of time in learning, only three studies have systematically reported data on amount of time spent in learning (Table 2.4). Two of these studies--Block and Jones--have been at the elementary level and one at the college level. Block and Jones both reported that mastery treatment required more time than non-mastery, a finding consistent with the theory that learning to mastery would require more time. The Merrill, Barton and Wood (1970) finding that the mastery group required less time than the non-mastery group is, on the surface, inconsistent with the assumptions of mastery learning.

While Block reported a significant difference in both achievement and retention between the highest criterion and control group, Jones did not. The difference in their results may be attributed to a difference in time. Block found there were no significant differences in time across the three higher criterion levels--both took about 84 minutes. However, this time was about 35 minutes more than the 49 minutes of the non-mastery group. This was a substantial increase. Jones likewise reported time spent in minutes, but the proportion of additional time spent was much less. Block reported 50% more time for mastery over non-mastery treatments; Jones reported only 14% more time. The fact that Block found that higher criterion mastery procedures significantly affected achievement and learning while Jones

Table 2.4

Learning Time Under Conditions of Mastery and Non-Mastery Learning

Study and Year	Number in Sample	Level and Content Area	Duration	Results
Block, J. H., 1970	91	8th, Algebra	5 days	75%, 85%, & 95% groups spent about same total amount of learning time on three units, about 84 minutes, compared to 66 minutes for 65% level and 49 minutes for control group.
Jones, F. G., 1974	20 classes	7th, Geography	4 weeks	Mastery treatment took significantly more time than non-mastery varying from 88 minutes for low aptitude to 60 minutes for high aptitude students.
Merrill, M., Barton, K., and Wood, L., 1970	40	College, Imaginary Science	5 lesson span	Time in learning decreased for experimental group with specific review procedure. Even though students received more material than control group they took less time to complete lessons.

did not may be attributed not only to differences in subject matter but to differences in time. It is possible that the procedures Jones utilized in his study did not require the mastery students to spend enough additional time to make a difference, in comparison with the non-mastery treatment. However, it should also be pointed out that the Jones study utilized a workbook in addition to the text for both the control and mastery groups, whereas Block merely used a text, without any exercises between study and the summative test. It is possible that mastery procedures, such as that of Block, in reality compare good teaching with deliberately constructed poor teaching situations. In such a case, the mastery procedures make a difference because they require more time. Where control procedures approximate quality instruction, without the provision for formative evaluation, such as used by Jones, it is less likely that the mastery procedure will make a difference. This results from the fact that the feedback-correction procedures of a so-called mastery treatment do not require enough additional time, when the control or non-mastery treatment utilizes practice exercises, as in a workbook.

Findings in relation to the use of additional time spent in learning are somewhat contradictory in the context of other types of studies. Thus both Dale (1972) and Pelletti (1973) found that the increased time spent in using the Forced Inferential Response Mode to construct responses from a data base did not contribute to an increment in

learning over that of learning from a narrative text. Fishburne (1971); on the other hand, did find that a programmed text in archeological methods brought about a significant difference in achievement over a narrative text. He attributed the difference to the fact that the programmed text took more time for study than the narrative text.

The time dimension in mastery learning is one that requires more attention. One difficulty, however, is that there is no uniform definition of mastery, and that mastery studies might be no more than comparisons of optimum with minimum quality instruction. As noted in the first chapter, the unit completion time variable was omitted from the present study because the time constraints of a fixed instructional period did not permit all students to complete their work.

Aptitude and Achievement Under Conditions of Mastery and Non-Mastery

Five studies, summarized in Table 2.5 explore the relationship of aptitude to mastery. In view of the essential hypothesis of mastery learning, that mastery procedures may reduce the negative affect of aptitude on learning, it is surprising that so little attention has been given to some measure of aptitude as a concomitant variable in mastery learning research. Use of aptitude as a blocking variable permits the investigator to classify students by aptitude level and ascertain if mastery treatment is more effective

Table 2.5

Aptitude in Relation to Mastery Learning

Study and Year	Number in Sample	Level and Content Area	Duration	Criterion Level	Results
Fagan, J., 1975	470 High, Middle, Low Vocabulary Levels	7th, Geography	4 weeks	85%	Aptitude Measure: Vocabulary, Iowa TBS, 1971 Forms 5 & 6. Vocabulary level significantly related to achievement and retention irrespective of treatment group. No difference in achievement or retention by treatment.
Glaser, G., 1968	N.S.	k-6, Math	3 years	85%	Aptitude Measure: N.S. High ability students mastered 5 times as many units as low ability students. Range of unit completion to mastery was 1 to 60 days. $A + .72$ correlation between previous knowledge and number of units covered over a 3 year period.

Table 2.5 (Continued)

Study and Year	Number in Sample	Level and Content Area	Duration	Criterion Level	Results
Jones, F. G., 1974.	20 classes High, Middle, and Low Vocabulary Levels	7th, Geography	4 weeks	85%	Aptitude Measure: Vocabulary, Iowa TBS, 1971, Forms 5 & 6. Vocabulary level significantly related to achievement and retention, irrespective of treatment except for middle aptitude mastery which achieved and retained significantly more than non-mastery.
Kim, H., 1969	273 above & below average	7th, Math	N.S.	80%	Aptitude Measure: N.S. 50% of experimental students with below-average I.Q. achieved 80% correct on summative test compared to only 8% of control group. 95% of experimental students with above-average IQ reached criterion compared to only 64% of control group.

Table 2.5 (Continued)

Study and Year	Number in Sample	Level and Content Area	Duration	Criterion Level	Results
Wyckoff, D. B., 1974	4 classes	6th, Anthropology	10 weeks	70%	Aptitude Measure: IQ, Otis-Lennon Reading level and IQ level significantly related to achievement, irrespective of treatment. Mastery treatment positively affected achievement of low reading level but not high reading level or any IQ level

than non-mastery treatment for low aptitude students or if it merely raises the level of achievement for all aptitude groups.

Vocabulary sub-tests of reading tests have been used as aptitude measures in two studies (Fagan, 1975; Jones, 1975) and reading scores were used in a third (Wyckoff, 1974). IQ test scores have also been utilized (Kim, 1969; Wyckoff, 1974). Reading scores correlate highly with IQ scores, since the group intelligence test is essentially a reading test (Ames, 1968). Reading scores correlate highly with achievement in social studies (Thomas, 1967; Dumbleton, 1973). Consequently, low scores on either a reading test or group intelligence test normally predict poor school performance on achievement (Ames & Walker, 1964). Mastery learning strategies, according to Bloom, are specifically needed to prevent this anticipated low performance.

The results of mastery research involving aptitude measures (Table 2.5) indicate little ground for optimism. Kim is the only one to give positive findings. The individualized programmed instruction of Glaser (1968) reported a correlation of .72 between previous knowledge and the number of units covered over a three year period. Moreover, highest ability students covered five times as many units as lowest ability students. The Glaser study is significant because of the large number of the experimental sample and the long time span--three years. The three other studies of

Fagan, Jones, and Wyckoff report no significant difference by treatment across aptitude, although there were significant differences in performance by aptitude group.

Aptitude was included in this study as a concomitant variable to ascertain if working toward different criterion levels would increase the affect of treatment and decrease the impact of aptitude on learning.

Mastery Learning and the Social Studies

Except for psychology at the college level, the social sciences have not been the subject of frequent mastery learning research. The popularity of psychology in mastery studies at the college level, excluded from this review but listed in Appendix G, is probably due to the fact that the subjects are readily available in intact classes to the psychology professors conducting the experiments.

Four studies of the social sciences, exclusive of college studies, are summarized in Table 2.6. The findings of the four studies have been consistent, indicating that aptitude, as measured by reading, IQ, or vocabulary, was a more potent factor in student learning than treatment. The social studies are dependent upon the ability of pupils to read to process information; a high level of mastery may therefore be unrealistic for low reading ability students in social studies, even under mastery treatment procedures. Deficiencies in reading are cumulative through the years, and mastery procedures applied during the course of a short

Table 2.6

Social Studies and Mastery Learning Research

Study and Year	Number in Sample	Level and Content Area	Duration	Results
Fagan, J., 1975	470	7th, Geography	4 weeks	Mastery procedures did not overcome learning problems of low aptitude students.
Gaines, G., 1971	28 classes	5th-8th, Anthropology	5 weeks	Carroll's interaction hypothesis between ability to understand instruction and quality of instruction not sustained. Important to control teacher variable.
Jones, F. G., 1974	20 classes	7th, Geography	4 weeks	Only middle aptitude mastery group performed significantly higher than non-mastery group on posttest and delayed posttest.
Wyckoff, D. B., 1974	4 classes	6th, Anthropology	10 weeks	No significant difference between mastery and non-mastery. Mastery had no differential effect upon aptitude or sex of students.

treatment are insufficient to make a corrective impact. However, the four social science studies included in the table have all been conducted in association with the Anthropology or Geography Curriculum Projects, University of Georgia. It may be that the mastery treatments have lacked sufficient corrective power. On the other hand, the conditions of instruction in the four studies for control as well as mastery groups attempted to provide "quality" instruction. The lack of difference in mastery as compared with non-mastery treatments may also result from the fact that the quality of instruction provided the control group did not sufficiently differ from that of the mastery group. A major limitation of these studies has been the time constraints of a fixed instructional period which has not permitted all students, especially lower ability students, to complete the treatment unit. Therefore, the lack of time for unit completion defeats the essence of mastery learning procedures.

Gaines (1971) tested the affect of mastery with fifth, sixth, seventh, and eighth grade students using anthropology materials. Gaines applied two learning strategies that were assumed to differ in quality of instruction in order to test the interaction of ability to understand instruction and quality of instruction. His mastery strategy involved the use of multiple choice test feedback. His control strategy utilized lesson practice in workbooks with self-correction. He assumed that his mastery strategy differed in quality from the workbook strategy. From his results it was not

possible to confirm the Carroll Model's hypothesized interaction of ability to understand with the quality of instruction. The differential effectiveness of teachers may well have obscured the main treatment effect. In addition, the actual difference between the test feedback and workbook feedback strategies may not have been sufficient to produce significant achievement differences.

Jones (1974) reported that achievement differences by aptitude of seventh grade students were not eliminated when self-instructional geography materials were used. Students of high aptitude scored significantly higher than middle and low aptitude students as did students of middle aptitude over students of low aptitude on learning and retention. There were no differences on the times-to-testing between any of the aptitude levels because instruction took place within a fixed time limit. Jones suggested that the lack of teacher monitoring in administering the review tests may have contributed to the poor performance of the low aptitude students. Low aptitude students require close personal supervision by the teacher, frequent feedback and learning success (Stuempfig & Maehr, 1970).

Wyckoff (1974) used anthropological content to compare the achievement of sixth graders under mastery and non-mastery conditions. The investigator considered treatment by IQ and reading level. The study did not produce evidence that mastery learning makes a significant difference in mean achievement between mastery and non-mastery groups. Highly

significant differences were reported in mean achievement between those students with low IQ scores and high IQ scores and high reading ability and low reading ability. Fagan (1975), who measured aptitude by a vocabulary test to derive high, middle, and low vocabulary levels, also reported that differences in achievement and retention by aptitude were highly significant but not treatment.

The social sciences are part of the required curriculum throughout the elementary grades. Normally, three years of social studies are required in high school. Continuing efforts to improve achievement in social studies are needed. This study attempted to utilize procedures which would make learning of a population geography unit, presented in the form of a data base and constructed responses, more effective.

Origins of the Present Study

This study developed out of an interest in the relationship of different criterion levels of performance to pupil achievement. In 1970 Block conducted a study of "The Effects of Various Levels of Performance on Selected Cognitive, Affective, and Time Variables," referred to previously in this review. In contrast to other mastery learning studies which typically selected some arbitrary criterion level and compared mastery with non-mastery treatments, Block used four different criterion levels and a control group. The criterion levels were 95, 85, 75, and 65%; no criterion level was

assigned the control group. Since the mastery procedures were the same for each of the experimental groups and the control group, except for the correction procedure which did not apply to the control group, the significant experimental variable was the various criterion levels. In addition to cognitive measures of achievement, Block also was interested in the relationship of certain affective outcomes, such as interest in and attitude toward mathematics, so that he might ascertain an optimum fit, if any, between criterion, achievement, and affective outcomes.

Description of the Block Study

The sample consisted of 91 eighth graders from four classes in a lower-middle class, suburban, elementary school. The subject was a three-unit sequence in matrix algebra, presented in the format of a programmed text for individual use. The principal reason for selecting matrix arithmetic was its sequential nature; each instructional segment built directly upon prior segments.

Sixteen students in each class were randomly assigned to mastery treatments--four per treatment by criterion level--and the remainder were assigned to a non-mastery treatment. The control and the 65 and 75% groups were placed in one room and the 85 and 95% groups in another. The use of rooms made it possible to group students who were expected to need comparable amounts of learning time. Mastery procedures consisted of the following: text study, formative test, restudy

in alternate form of text, second formative test, teacher or peer tutoring, and third formative test. Students moved to the next unit when criterion level was attained on the formative test. If the criterion level was not attained after the third formative test, the student proceeded to the next assignment. The alternate form of the programmed text was a more detailed and complete explanation of the learning tasks. Review of the alternate programmed text was facilitated by a coded review sheet, which referred the student to the unlearned or misunderstood material. The duration of the study was five days, including pre-treatment and post-treatment administration. Four of the periods were 80 minutes, and one period was a 40 minute session for students who needed more time.

Cognitive data were collected by means of investigator constructed tests of achievement, transfer, and retention.

Retention was measured by an alternative form of the achievement test, administered as a delayed posttest two weeks after the achievement test. The retention test was preceded by a 15 minute review.

Results and Discussion

Results will be discussed in terms of cognitive and affective measures. There was significant difference in achievement in the 85 and 95% groups over the control group, but no significant difference in the performance of the 65 and 75% groups over the control group. The only mastery

group that showed any transfer advantage over the control group was the 95% group. The results of the retention test were the same as for the achievement test, significant differences over the control group at the 85 and 95% levels but not at the 65 and 75% levels. In terms of learning time, the mastery procedures were effective in requiring students at the higher criterion levels to spend more time in review. The five groups, experimental and control, spent about the same time in initial study of the text. The three criterion levels of 75, 85, and 95% required approximately the same amount of review time--31 minutes--so that total learning time for these groups was about 84 minutes compared to 49 minutes for the control group. By unit III, however, it would appear that mastery at the 95% level contributed to increased learning efficiency. Students in this group took the same text time as the control group, but on the first formative test made a percentage score of 74 compared to 54 by the control group.

Affective results indicated that attainment of either the 85 or 95% levels yielded significantly greater interest in and attitude toward the arithmetic than the non-mastery treatment as measured concurrently with achievement. Only attainment of the 85% level yielded significantly higher scores than the non-mastery treatment as measured concurrently with retention. Both interest in and attitude toward arithmetic scores decreased from the short-term to long-term

administration, but only the 85% level produced sustained high levels of interest and attitude over time.

Block's findings indicate that by varying the performance level required of students and by using simple feedback-correction techniques to ensure that the prescribed level is attained, end-of-sequence achievement can be increased around a high average score. He suggests that 80 to 85% correct is a more realistic standard to maintain throughout a mastery learning sequence. Setting standards too high may be wasteful in terms of teacher and student time, and may have a negative effect on student motivation. The lower standard tends to provide more opportunities for student success and thus increases the amount of positive reinforcement.

The limited generalizability of Block's findings was the main reason for the present study. The investigator wished to extend the format of the Block study using a larger sample, a longer learning sequence, and different subject matter. In addition, aptitude, as measured by a vocabulary test, was included as a second independent variable, to ascertain whether the maintenance of any particular criterion level increased achievement of lower aptitude students on a summative posttest measure.

The characteristics of the present study and the Block study (1970) are compared by variables in Table 2.7. The most striking differences in the two studies are in the length, of 15 days compared with five; the subject matter, population geography compared with matrix algebra; sample size, 734

Table 2.7

Variables Used by Block (1970) and Contreras (1975)

Variable	Block Study	Contreras Study
SAMPLE		
Number Subjects	91	734
Number Classes	4	24
Subject Loss	15	33
% Subject Loss	16%	48
No. of Teachers	4	5
GRADE LEVEL	8	7
DURATION OF STUDY		
Inclusive Time	5 days	15 days
Treatment Time	3 days	14 days
Retention Elapsed Time	2 weeks with 15 minute review	3 weeks, no review
INSTRUCTIONAL MATERIAL		
Subject Matter	Matrix Algebra	Population Geography
Nature of Material	Sequential and hierarchical; skills build on previous learning	Factual, using descriptive, analytical, and synthetic statements and generalization; sequencing logical but not hierarchical

Table 2.7 (Continued)

Variable	Block Study	Contreras Study
Format of Material	Individual tutor test in alternate forms	Individual self study text using constructed responses on incomplete sentences drawing on a database (FIRM)
DATA ANALYSIS Concomitant Variable	None	Vocabulary knowledge, as measured by Iowa TBS, Forms 5 and 6, 1971
Design Unit of Analysis	Pre-test, posttest Individual pupil	Post-test only Aptitude group within class by treatment
Principal Method of Analysis	Analysis of Variance	Analysis of Variance
MEASURES Cognitive	Achievement, end of unit Retention, after two weeks Transfer	Achievement, end of unit Retention, after three weeks ---

Table 2.7 (Continued)

Variable	Block Study	Contreras Study
Time	Time to Criterion	---
Affective	Interest in Algebra Attitude Toward Algebra and Its Difficulty, Huson, 1967	--- Attitude Toward Any Sub- ject, Short Form, Remmers, 1960
MASTERY PROCEDURES First Treatment Feedback Second Treatment Feedback Third Treatment Feedback Proceed	Tutor Text Formative Test Alternate Form Tutor Text Second Formative Test Tutoring Third Formative Test To next lesson, without meeting criterion	FIRM Text Formative Exercise Review of same data base in FIRM Same formative exercise --- --- To next lesson, without meeting criterion

compared with 91; formative tests, two compared with three; the unit of analysis, the group within class compared to the individual; and concomitant variable, vocabulary knowledge, compared with none. While other mastery studies in social science have compared mastery with non-mastery procedures and have also involved aptitude as a concomitant measure, the distinctive characteristics of the present study are the use of differential criterion levels and aptitude applied to mastery learning in a social studies unit. Thus the main objective of the study is to ascertain if differential criterion levels interact with mastery and aptitude to produce significant differences in achievement, retention, and attitude toward the subject. The inclusion of aptitude as an independent variable is important because a major premise of mastery learning is that a feedback-correction component included throughout a learning sequence will enable all students to master what the school assigns to a high level of attainment.

The next chapter will review the general methodologies and specific procedures used in testing the hypotheses raised in Chapter I and enlarged upon in the review of the literature.

CHAPTER III

PROCEDURES AND METHODOLOGIES

This chapter is divided into six parts. The first three describe treatment, data collection, and contextual variables. The next two parts describe experimental design and statistical procedures employed. The final part describes the limitations of the research.

Description of Treatment

This section describes seven aspects of the study: treatment variable, instructional unit, sample selection, duration of the study, procedures for the pilot study, orientation of teachers, and procedures for the final experimental study.

Treatment Variable

The treatment variable in this study, as noted in Chapter I and II, consisted of the differential affect of three criterion levels of achievement on final achievement as measured by the summative test. Customarily, performance on an achievement test is a dependent variable because it is an outcome of some treatment variable. In this mastery learning study, however, the requirement that various groups attain different

criterion levels to ascertain the differential impact of performance levels, makes what is traditionally the dependent variable--achievement--the experimental variable. The three treatment levels, to which intact classes were randomly assigned were: T_1 , 90%; T_2 , 80%; and T_3 , 70%.

The same unit of instruction and classroom procedures were utilized by all treatment groups, irrespective of criterion level.

Instructional Unit

The instructional unit was Population Growth in the United States and Mexico (Dale and Rice, 1972), written for the Geography Curriculum Project of the University of Georgia. The text presents basic concepts of demography and population geography in a historical and comparative format. The unit materials consist of a pupil text and a pupil answer booklet and may be used as an individualized tutor text, as a text for class-paced instruction, or a combination of individualized and class paced instruction. Appendix A contains the text's table of contents and a sample lesson from each of the three parts within the unit.

There were four principal reasons for selecting the unit Population Growth in the United States and Mexico. The first reason for selecting the unit was the structured nature of the material, with clearly designated learning outcomes and feedback procedures. According to Carroll (1963) and Bloom (1968), structured material of this type lends itself to mastery

learning procedures. The instructional unit is organized according to a method of presentation called the Forced Inferential Response Mode (FIRM). It uses incomplete sentence stems to force students to derive information from a data base in order to construct a series of sequential responses. The correctly completed stems compose a logical narrative of information contained in the data base. In this unit, the data base consisted of maps, graphs, tables, and charts. The FIRM text is accompanied by a response booklet which permits the student to obtain feedback as to the correct nature of his response.

A second reason for selecting the unit Population Growth in the United States and Mexico was the appropriateness of readability. Readability of the unit was not reported by Dale, even though it was written for students in the middle grades. It had previously been tested with fifth, sixth, and seventh grade students.

Two standardized reading formulas were used by this researcher to determine readability. The Rudolf Flesch (1949) formula for readability was one formula applied. Four reading ease scores were computed from sample selections taken from the beginning, the middle, and the end of the unit text. These scores indicated that the sample selections had readability levels of grades five, six, and seven (Table 3.1).

A second standardized reading formula used to determine readability was the Fry (1965) formula. The mean number of sentences per 100 words was 7.2, and the mean number of

Table 3.1
Computed Reading Ease Scores According to the
Rudolf Flesch Readability Formula

Page No.	Reading Ease Score	Grade Level
8	92	5
48	78	7
54	81	6
80	73	7

syllables per 100 words was 138. On the graph for estimating readability, the intersection of these two means indicate a seventh grade reading level (Table 3.2).

Third, there was the appropriateness of subject matter. Seventh grade social-studies curriculum in Texas, where the research was conducted, deals with Latin American and Texas history. Mexico is often a topic in much of the state's social studies curriculum because of its proximity to the United States and the resulting cultural diffusion between the two countries.

Fourth, and more important, was the specific relevance of the unit to the geographical location of El Paso, Texas, the city where the research was conducted, and to the large Mexican American population of the area. Together Ciudad Juarez and El Paso form a border metropolitan area of approximately one million people. The majority of the El Paso population is of Spanish surname (El Paso Chamber of Commerce, 1974). The unit explained, to a large extent, the historical development of the Mexican American population now concentrated in the Southwest. It also reinforced previously taught material on Texas and Mexican history. The decision to use this unit was made after a sample had been identified for the research and formal arrangements with an El Paso school district had been finalized. The procedural arrangements with the El Paso School District are given in Appendix H.

Table 3.2
 Computed Reading Ease Score According to
 Fry Readability Formula

	Sentences per 100 words	Syllables per 100 words
100-word sample Page 4	9.1	122
100-word sample Page 32	8.8	148
100-word sample Page 82	6.6	130
	<u>3) 21.7</u>	<u>3) 414</u>
Average*	7.2	138

*Readability measured at grade seven level.

Sample Selection

This researcher made arrangements with the officials of the El Paso Independent School District in Texas to obtain 24 seventh grade classes in five schools for the experimental study. Teacher selection was based upon a willingness to take part in the experiment.

Duration of the Study

Preparations for conducting a pilot study prior to the experimental study were made by the researcher for a 15 day instructional period from February 3, 1975 to February 21, 1975.

The experimental study was conducted over a 15 day instructional period from March 3, 1975 to March 21, 1975. During this period the three criterion treatment groups studied the unit Population Growth in the United States and Mexico. At the end of the 15 day instructional period, a posttest of geography achievement and an attitude scale were administered. A delayed posttest of geography achievement was administered three weeks after treatment to measure retention. The time schedule of the study is included in Appendix I.

The procedures employed in the pilot and experimental study are described in the following sections.

Pilot Study

A pilot study was conducted from February 5, to February 21 to monitor treatment procedures to be used in the experiment. Three classes from the El Paso Independent School District were utilized.

In the pilot study classes were randomly assigned to the 90, 80, or 70% criterion levels. The prescribed criterion level was to be attained before proceeding from one lesson to the next. All classes received the same unit text and answer booklet. Students also received a personal score log which included the minimum number of correct lesson responses needed to meet the prescribed criterion level. A sample of the score log is given in Appendix B.

Several important procedural changes resulted during the pilot study itself. One change was the elimination of a semi-class-paced mastery strategy. Originally, the procedure called for a discussion by the teacher of a lesson's data base before students proceeded to study the lesson on their own and complete the unfinished sentence stems of the respective lesson.

The teacher, therefore, was to play an important role in the study by systematically introducing to the class each data base of the unit's 41 lessons. However, this semi-class-paced procedure was abandoned a few days after the pilot study got underway. By the third day few students were on the same lesson. The teacher's explanation of a

given data base created a disturbance to those who were behind or beyond the data base under discussion.

It was decided to allow students to continue through the unit on an individual basis, with teacher assistance to the individual. The students were instructed to request teacher assistance whenever difficulty was encountered. In this manner, explanations, clarifications, and other types of guidance were offered to children on an individual basis and as the need for help arose.

A second change which occurred during the pilot study was the elimination of study groups. Originally, study groups were designated as one means of providing feedback-correction to those students who had failed to achieve the prescribed criterion. After the first day of the pilot study, it was evident that this procedure was difficult to implement. Instead, class monitors, those who had completed the unit or were far ahead of the class as a whole, were used to assist students having learning problems. Monitors were instructed not to give answers to students, but to help them to derive the answers from the data base themselves.

All subjects, irrespective of the criterion level required to reach, followed the steps described below.

1. The date the student began the treatment was recorded in the personal score log.
2. The treatment began with lesson one. Students studied the data base.
3. The incomplete sentence stems, corresponding to each lesson's data base, were finished. The responses were written in the unit text.

4. When all sentence stems were completed, the finished lesson was shown to the teacher. The teacher gave the student an answer booklet to check his responses.
5. The number of correct responses was recorded in the pupil's personal score log under the column "First Exercise" (Appendix B).
6. If the student answered at least the minimum number of stems correctly, as indicated on his score log for each lesson, he proceeded to the next lesson.
7. If the student did not achieve criterion on the lesson, he reviewed the data base of the same lesson a second time. He studied the data base until he thought he understood it well enough to repeat the written exercise.
8. A sheet of paper was obtained from the teacher in which the answers were written.
9. The correction procedure was repeated as indicated in step 4. This score was entered in the log under the column "Second Exercise." Whether or not criterion had been attained, students proceeded to the next lesson. The pilot study had shown that almost all students achieved criterion on the repeated exercise.
10. Students proceeded at their own pace until the unit was completed. Students recorded the date of unit completion in their score logs.
11. The final posttest was administered on the 15th day of treatment.

Throughout the learning unit, the teacher and appointed monitors assisted students requesting help in understanding a lesson. If the unit was worked upon at home, students were required to record the time spent studying.

The researcher conducted the pilot study herself for two reasons. First, the original classroom teacher did not have sufficient time to study the unit or to review the

mastery procedures prior to the beginning of the pilot study. In order to assist students with learning difficulties, a thorough understanding of the unit was essential. Secondly, a participant-observer approach enabled the researcher to retain or modify mastery procedures for the final experimental study. Having directly experienced the pilot study facilitated the teacher orientation meeting between the researcher and the teachers in the final experimental study.

Orientation of Teachers

A three hour orientation meeting was held with the five participating teachers of the five cooperating schools on February 27, 1975 four days prior to the beginning of the treatment. February 27 and 28 had been designated as in-service days for teachers of the El Paso Independent School District, and principals of the participating teachers granted them permission to attend the orientation arranged by the researcher. Teachers had previously received copies of the text Population Growth in the United States and Mexico and had reviewed the unit prior to orientation.

The orientation stressed the importance of the teacher in the experimental study. The unit required close teacher supervision and participation. It was important that they be completely familiar with the unit in order to assist students encountering difficulties with the lessons. Therefore, teachers were asked to complete the unit themselves before the beginning of the experimental study. Some teachers,

however, did not complete the unit until the following week.

On the first day of the treatment, students were to receive an explanation on the use of maps, graphs, charts, and tables to facilitate completion of the information statements. The researcher demonstrated this lesson to the teachers and provided them with additional aids to facilitate the teaching of this introductory lesson. The directions to the students and teacher are given in Appendix C. The sample questions for at least one of the lessons involving a map, a chart, a graph, and a table are given in Appendix D. This lesson format for class paced mastery was used by teachers to introduce new material, but was not used as originally intended with each lesson.

Experimental Study

Except for one major change pertaining to the availability of answer booklets, the experimental procedure resembled the process pursued in the pilot phase of the study. In the final study answer booklets were not distributed to individual students. Individuals were required to show their completed lessons to the teacher, who in turn released an answer booklet to them for self correction in a specially designated area of the classroom.

This alternate measure was thought to reduce the probability of copying answers without study of the data base. The researcher had noted in the pilot study that some

students tended to locate responses in the answer booklet before devoting ample time to the study of a data base to infer an answer. The reluctance to persevere in the learning task was thought to be enhanced by the immediate availability of an answer booklet which each individual could readily resort to. Therefore, teachers of the final study decided to retain control of answer booklets. However, there is some doubt that students reached the stipulated criterion level because students self-reported their individual progress in personal score log. Even though students were required to show their completed lessons to the teacher before obtaining an answer booklet for checking purposes, the teacher did not have sufficient time to systematically check the pupils' exercises for accurate reporting.

The next part on data collection describes the two achievement tests and the attitude scale administered to students.

Description of Data Collection

This section describes the word meaning test, the achievement test, and the attitude scale used in the research.

Word Meaning Test

Data for the treatment by blocks design was provided by the word meaning section of the Iowa Test of Basic Skills: Form 6, Level 13, a 48-item, four-foil multiple choice test (Lindquist & Hieronymus, 1971). It was administered prior

to treatment to all students in the 24 participating classes. There is evidence to indicate that vocabulary achievement correlates highly with achievement when the University of Georgia Anthropology or Geography Curriculum Project materials are used (Thomas, 1967; Gaines, 1971; Jones, 1974).

Word meaning knowledge and total reading achievement are also highly correlated. Vocabulary knowledge is essential to reading ability, and has been used to predict success in school (Traxler, 1945; Johns, 1972; Duffy, et al., 1972). Therefore, a vocabulary score was utilized as the blocking variable because of its high correlation with total reading comprehension and achievement and economy of time in administration.

The word meaning section of the Iowa Tests was also chosen for its high test reliability. According to the Technical Manual (Lindquist & Hieronymus, 1971), the grade seven vocabulary test has a test reliability of .89, and the reading test has a reliability of .92. The standard error of measurement for the vocabulary test is 3.0. Word meaning had a correlation of .81 with reading comprehension, the highest correlation of any sub-test of the Iowa Battery.

The researcher utilized the ANLIGH (Analysis of Item and Test Homogeneity) program to compute the reliability of the Word Meaning Section of the Iowa Tests of Basic Skills: Form 6, Level 13 with a random sample of 100 students from the population of 734 seventh grade students from the El Paso Independent School District who participated in this study.

Computation of reliability by Cronbach's Alpha Indices gave a coefficient of .83.

Achievement

The summative test used in this study was a 50-item, four-foil multiple choice test developed by Dale in 1972. The content validity of the test, according to Dale (1972, p. 52), was controlled by a table of specifications and comparisons of content and test items. He reported a reliability, Kuder-Richardson #21, of .91. Because the population of the present study was geographically and ethnically different from that of the Dale study, the investigator computed a new reliability score for the subjects in this mastery study. The ANLITH program gave a coefficient of .93 (Cronbach's Alpha Indices). A copy of the test is given in Appendix E.

Attitude Scale

It was evident to the researcher that there was no standardized scale in existence that measured attitude toward mastery learning with geographic subject matter. However, the volume Scales for the Measurement of Attitude by Shaw and Wright (1967) reports 172 attitude scales with at least minimal reliability and validity. The authors suggest that "Attitude research has been hindered by the inaccessibility of existing attitude scales resulting in less-than-optimum advances in the scientific analysis of attitudes." Therefore, they reported scales which they considered to be among the most useful in meeting current research needs.

Among the many generalized scales contained in the Shaw and Wright volume was one scale that measured attitude toward any school subject, developed in two equivalent forms by Silance and Remmers (1934). Silance and Remmers (1934) reported equivalent-forms reliabilities ranging from .81 to .90, using different school subjects as attitudinal referents. Ferguson (1952) cites the following reliabilities for four different subjects: biology, .81 ($n = 269$); chemistry, .70 ($n = 771$); English, .68 ($n = 705$); mathematics, .74 ($n = 579$). Stoddard (1975) reported a reliability of .92 ($n = 30$ classes) for American history.

Bolton (1938) validated the scale for construct validity by using criterion groups measured for interests and values in the area of mathematics. Strunk (1957) offered some evidence of concurrent validity with a correlation of .39 between this scale and the scores of 130 subjects on a graphic rating scale of expression of interest in a psychology course.

This scale seemed reasonably valid and reliable for purposes of this study; however, there was also a shortened form of the 45-item scale developed by Remmers in 1960. The investigator selected the Remmer's 17-item short A Form rather than the long 45-item A Form for the study because the short form would take less time for administration. Time was essential to consider in this study because students were administered two posttests on the final day of the treatment, and the investigator thought that there might be a lack of time for the completion of both tests by all

students in one class period. Both posttests--the geography achievement posttest and the attitude scale--were successfully completed by the subjects of the pilot study when the short form of the attitude scale was administered.

According to a review of the Remmer's shortened test forms by Campbell (1953), as reported in The Fourth Mental Measurements Yearbook, the test manual gives no criterion for the selection of the retained items. No correlations between the original and short forms of the various generalized scales developed by Remmers and his associates (The Fourth Mental Measurements Yearbook, 1953) are provided, nor is data on reliability reported. Evidence of validity comes from the older forms of the scales, that is, the typically high correlations between the generalized scales and Thurstone scales on similar attitudes (Campbell, 1953 & Clark, 1953). Both reviewers of the Remmer's shortened test forms, Campbell (1953) and Clark (1953), note the unique usefulness of the Attitude Toward Any School Subject scale as a generalized scale and predict their continued use.

Although no data on reliability or validity for the short form was found, the investigator assumed that reliability and validity would differ only as a result of the decrease in items. The present investigator determined reliability of Form A by the equivalent forms method. Form B of the 17-item Attitude Toward Any Subject scale was administered to a sample of 111 students in the study after

a delay of three weeks and on the same day the delayed post-test was administered. The responses of these students on Form B were compared with their responses on Form A, giving a correlation of .93. On the basis of the correlation of short Form A and B in this study, Form A is considered a reliable measure of the possible influence of different criterion levels on attitudes for this study.

A sample of Remmers Short Form A, modified as "Attitude Toward A Geography Unit," is included in Appendix F.

Description of Contextual Variables

This section discusses those contextual variables which may have affected the treatment but could not be controlled in the experiment. They include the effect of the community and the school district, the school, and the teachers, variables considered pertinent to the experiment's external validity. The results of this study can be generalized to groups using similar materials and with characteristics similar to those in the treatment sample.

Community and School District

The study was conducted in the El Paso Independent School District of El Paso, Texas. The population of the city of El Paso is approximately 360,725. For the county of El Paso, the population numbers some 400,971 people. The proportion of racial categories for the county is as follows: Caucasian 96.1%, Black 2.8%, Indian .2%, Japanese .2%, and

all other .8%. Persons of Spanish language or surname comprise 58.1% of the city's population and 56.9% of the county's population (El Paso Chamber of Commerce, 1974).

El Paso is adjacent to the Rio Grande River and Ciudad Juarez, Mexico. Together these two form an economically integrated metropolitan area with a population of approximately 877,347 people. International border crossings from Juarez, Mexico to El Paso number 41,936,873 in 1973 (El Paso Chamber of Commerce, 1974).

The economic base of the city and county of El Paso is apparel manufacturing, smelting and refining of metals, and petroleum refining and the distribution of natural gas. El Paso's economy also includes major military expenditures from Ft. Bliss, a major national defense training center.

Twenty-one high schools, five junior high schools, and 91 elementary schools serve the city's 108,084 students (El Paso Chamber of Commerce, 1974). Seven small independent school districts outside the city limits serve an additional 7,744 students. The El Paso Independent School District consists of eight high schools, 47 elementary and intermediate schools, three intermediate schools, and one junior high school.

Bussing is not used in the county of El Paso to facilitate a balance of Mexican American and Anglo American students in every school. The director of Pupil Services for the school district (C. F. Hart, Personal interview, March 21, 1975) reported that bussing was rejected in El Paso

because the city itself is integrated. To maintain this balance, the El Paso Independent School District is subdivided into small divisions. Students are required to attend those schools within the sub-districts in which they reside. Two of the high schools, however, serve almost 100% of students of Spanish surnames. Some 11 elementary-intermediate and junior high schools serve as feeder schools for these two high schools.

Characteristics of the Schools in the Study

The 24 classes that participated in this study were located in five schools in the El Paso Independent School District. Each teacher completed a questionnaire which described both teacher and school characteristics considered relevant to the study. The reported data relating to the socio-economic status of pupils in this study is based on informal teacher use and not on precise SES data. School characteristics are summarized in Table 3.3.

School A. This school was constructed in 1967. It has 54 teachers and is administered by an appointed principal. It is an elementary-intermediate school, containing grades k-8.

The classes are departmentalized, and students are heterogeneously grouped. This composition of the student body is approximately 50% Anglo American and 50% Mexican American. Socio-economically, the surrounding school area ranges from upper class to lower class. The school population

Table 3.3
School Characteristics Summary Table

School	Date Con- structed	No. of Teachers	Grades	Percent Ethnic Composition		SES			Dominant Language
				Anglo	Mex. Am.	U	M	L	
A	1967	54	k-8	50	50	25	50	25	English
B	1958	58	6-8	1	98		66	34	Spanish
C	1972	11	6-8	60	40	25	60	15	English
D	1958	47	k-8	50	50	25	50	25	English
E	1925	47	7-9		100		25	75	Spanish

00105

is approximately 50% middle class, 25% lower class, and 25% upper class. Students from two extremes attend this school-- children from farm workers and children from the community's wealthiest sector.

School B. This school was constructed in 1958. The school has 58 teachers and is administered by an appointed principal. This intermediate school contains grades six, seven and eight. Classes are departmentalized and heterogeneously grouped.

Approximately 98% of the school population is Mexican American, 1% is Black, and 1% Anglo American. About one third of the population is from a lower class background, living primarily in government housing. Another third of the school is of a lower middle class status, and the remaining third of the population is classified as middle middle class. For the majority of the students, Spanish is the dominant language.

School C. This school was constructed in 1972. The school has 11 teachers and is administered by an appointed principal. This intermediate school includes grades six, seven, and eight. Classes are departmentalized and heterogeneously grouped.

Like school A, this school is also in a socio-economically mixed area. About 25% of the children come from an upper class background. Some 60% of the school population is classified as middle class, and about 15% is of a lower class status, living in governmental housing.

The influence of the United States Army is strong in this geographic area which houses a large number of army related people.

School D. This school was constructed in 1958. The school has 47 teachers and is administered by an appointed principal. This elementary-intermediate school includes grades k-8. Classes are departmentalized and classes are heterogeneously grouped.

The socio-economic status of the area varies from middle class to lower class. Approximately 50% of the school population is Anglo-American and the remaining 50% Mexican American.

School E. This school was constructed over 50 years ago. Originally this was one of the city's oldest high schools. Two years ago it became an intermediate school. The school has 47 teachers and is administered by an appointed principal. This intermediate school includes grades seven, eight, and nine. Classes are departmentalized and heterogeneously grouped.

The school population is almost 100% Mexican-American. Most of the students come from a low socio-economic background. Spanish is the dominant language for most, if not all, of the school population. Government housing is concentrated in this school's geographic location.

Characteristics of the Teachers in the Study

Five grade seven teachers from the El Paso Independent School District participated in this study. The biographic teacher characteristics data in this section was collected from each teacher by the use of a questionnaire and is summarized in Table 3.4. The judgments of teacher behavior by the investigator are based on frequent but informal observations of teacher and pupil classroom interactions.

Teacher A. This teacher was 31 years old, male, and had six years teaching experience. He held a Bachelor of Arts degree in history and an M. Ed., plus 30 semester hours beyond the M. Ed. He reported that he had completed 6 semester hours in geography. This teacher taught seven sections of seventh grade social studies.

This teacher had excellent discipline in his classroom. Both teacher and student expectations appeared to be clearly defined and understood by each side. The teacher maintained good rapport with the students and interacted with them very frequently throughout the experimental study. A very pleasant classroom atmosphere prevailed whenever the researcher observed this teacher and his classes.

Teacher B. This teacher was also 31 years old, male, and had four years teaching experience. He held a Bachelor of Science degree in Education with a major in secondary education. He had completed nine semester hours in geography. This teacher taught six sections of seventh grade social studies.

Table 3.4
Teacher Characteristics Summary Table

Teacher	Age	Sex	Years Teaching Experience	Degree	Semester Hours in Geography	No. of Classes Taught	Classroom	Average Number of Pupils per Class
A	31	M	6	B.A. (History) M.Ed.	6	7	Control oriented	32
B	31	M	4	B.S.Ed.	9	6	Control oriented	28
C	30	F	5	B.S.Ed.	3	3	Freedom oriented	29
D	33	F	1	B.S.Ed.	3	6	Control oriented	34
E	25	F	4	B.S.Ed.	3	2	Control oriented	23

This teacher also had excellent classroom discipline. Students were always observed in their seats and appeared to be working through the unit the entire period. This teacher's classes were generally very quiet and with less student-teacher interaction taking place than with teacher A.

Teacher C. This teacher was 30 years of age, female, and had five years teaching experience. She held a Bachelor of Arts Degree in Elementary Education. She reported that she had completed three semester hours in geography. She taught three sections of seventh grade social studies.

Discipline in this classroom appeared to be somewhat more lax than the class mentioned above. A considerable number of informal student to student interactions were evidenced by the researcher. However, this teacher very frequently worked with individual students and their problems with the unit text.

Teacher D. This teacher was 33 years of age, female, and had one year of teaching experience. She held a Bachelor of Science degree in Elementary Education. She reported that she had completed nine semester hours of geography. She taught six sections of seventh grade social studies.

This teacher averaged the largest number of students per class, approximately 34 pupils. Conditions in her classroom appeared somewhat crowded in comparison with any of the other classrooms. Classroom discipline appeared satisfactory in each of this teacher's classes. This teacher

constantly interacted with her students throughout the experimental study and appeared to have developed good rapport with her pupils.

Teacher E. This teacher was 25 years old, female, and had four years teaching experience. She held a Bachelor of Science degree in Secondary Education. She reported that she had completed three semester hours in geography. She taught five classes of social studies.

The teacher averaged 23 students per class, the lowest number of all other teachers. She maintained very good discipline and was able to keep close track of each pupil's progress. She also interacted with students frequently as they proceeded through the unit text. Students remained in their seats while she generally worked her way from one pupil to another.

Summary of Contextual Variables

The five participating schools were similar in organization, administration, and plant facilities. The schools differed, however, in student population. The population of two schools was predominantly Mexican American. In one, socio-economic status was approximately 75% lower and 25% middle class. Students of the second school were also from a lower or middle class background. The dominant language for both school populations is Spanish. This means that children generally speak English in the classroom and with their teachers where subject matter is concerned. However,

students speak Spanish to one another in most informal interactions both inside and outside of the classroom. The home language is Spanish.

In the three remaining schools, the Mexican Americans are a minority. Anglo students comprise about 50% of the population in each of the three schools. Most interactions between an Anglo American and a Mexican American child require that the latter speak English rather than Spanish. Most Anglo children in these schools do not speak Spanish fluently enough to converse in the language. Mexican American students in these schools tend to be more bicultural than those students in the predominantly Mexican American schools.

Eight classes from predominantly Mexican American schools participated in the experimental study. The remaining 16 classes were from mixed schools where approximately one half of the student body was Mexican American and the other half Anglo. When class means were rank ordered for all 24 classes according to the word meaning section of the Iowa Tests of Basic Skills: Form 6, Level 13, the classes from the two predominantly Mexican American schools occupied the bottom eight positions, or bottom third, of the range. The mean of this bottom third was 16.85 on the word meaning test, a grade equivalent of 5.5, while the mean of the other 16 classes was 21.99, a grade equivalent of 6.7. Therefore, the mean vocabulary equivalent of the predominantly Mexican

American schools was slightly over a year behind the predominantly Anglo schools (Table 3.5).

The observed differences between the three treatment groups regarding the personal attributes of the teachers and individual school characteristics were deemed to be minor. Because of the random assignment of classes to treatment, classes from the predominantly Mexican American schools were distributed across treatments. Thus, the investigator concluded that there were no contextual variables, other than treatment, that accounted for observed differences between the three treatments on the posttest. Table 3.6 shows the distribution of classes across treatments by ethnic groups.

The next two parts of this chapter relate the technical aspects of the experimental research.

Experimental Design

A 3 x 3 fully crossed analysis of variance (ANOVA) for each criterion measure was employed with the posttest data of this study. This design is shown in Table 3.7. The same design was used for all three dependent variables. To avoid needless repetition of the design, only one experimental layout is shown with achievement as the dependent variable.

Random Assignment of Classes to Treatment Groups

There were two steps in the randomization process. First, the 24 classes were randomly assigned to one of three

Table 3.5
Vocabulary Levels Across Two Types of Schools

Type of School	No. of Schools	No. of Classes	Class Means on Word Meaning Test	Grade Equivalent
Predominantly Mexican American Population	2	8	16.85	5.5
Mixed Population 50% Anglo American 50% Mexican American	3	16	21.99	6.7

Table 3.6
Distribution of Classes Across Treatment
by Ethnic Groups

Ethnic Group	90% Criterion Group	80% Criterion Group	70% Criterion Group
Predominantly Mexican American	3	2	2
Mixed Mexican American and Anglo American	5	6	5

Table 3.7
Experimental Layout (ANOVA)

	T ₁ -90%	T ₂ -80%	T ₃ -70%	
	C ₁₋₈	C ₉₋₁₆	C ₁₇₋₂₄	
A ₁	Y _{.111}	Y _{.121}	Y _{.131}	Y _{.1.1}
A ₂	Y _{.211}	Y _{.221}	Y _{.231}	Y _{.2.1}
A ₃	Y _{.311}	Y _{.321}	Y _{.331}	Y _{.3.1}

Y_{..11}Y_{..21}Y_{..31}

Independent Variables:

Treatment
Vocabulary Levels

Dependent Variable:

Achievement

T represents treatments: $t = 3$

A represents vocabulary levels: $a = 3$

C represents classes: $c = 24$

Subscript Order: classes, vocabulary levels,
treatment, measure

treatment groups. Treatment was then randomly assigned to the groups. The first treatment group was to achieve to the 90% criterion level; the second treatment group to the 80% criterion level; and the third group to the 70% criterion level.

A treatment by blocks design was adopted for examining the effect of mastery learning on students of varying aptitudes, especially lower ability students. Aptitude, as measured by a vocabulary test, was selected as the blocking variable. The reason for using a treatment by blocks design is related to the premises of mastery learning, discussed briefly in the introductory chapter and the review of the literature. An explicit premise of mastery learning is that feedback-corrective procedures are instructional techniques which raise the level of achievement of lower performing pupils in traditional instruction. This premise, as previously noted, tends to emphasize the factors of the environment, over what happens in the instructional situation, and to ignore the characteristics a learner brings to instruction. The psychological literature has repeatedly demonstrated that pupil aptitude rather than the type of treatment is more important in cognitive achievement (Duchastel & Merrill, 1973; Oswald & Fletcher, 1970; Dalis, 1970; Merrill & Towle, 1972; Duffy, 1972). The treatment by blocks design permits the researcher in mastery learning to ascertain if mastery procedures eliminate learning for the lower

aptitude student, when aptitude is measured by a vocabulary test, as in this study.

Students were administered the word meaning section of the Iowa Tests of Basic Skills: Form 6, Level 13 on the first day of the treatment. A subsequent analysis of vocabulary achievement across criterion level of treatment revealed a homogeneity of group means (Table 3.8). The subjects were rank ordered by class on the basis of the concomitant score and then divided into three blocks: a high, middle, or low block. The designated ranges permitted a minimum of three individual scores per classroom to fall within each vocabulary level, except for the high vocabulary group in one classroom which contained only two scores (Table 3.9). A discussion of how the vocabulary ranges were determined follows.

Distribution of Students by Treatment and Vocabulary Level

Five teachers and a total of 24 seventh grade classes with 734 students participated in the study. The number of students within classes were classified by vocabulary level and criterion level in Table 3.9. Not all 734 students were used in the data analysis. Thirty-four students were omitted from the data analysis for one of three reasons: (1) prolonged absences from school, (2) withdrawal from school prior to the completion of the study, and (3) late enrollment into school after the study was almost completed.

Table 3.8
Mean Vocabulary Achievement Scores
by Treatment Groups

	T ₁ 90% Criterion	T ₂ 80% Criterion	T ₃ 70% Criterion
Vocabulary Means	21.35	21.66	21.79

Table 3.9

Number of Students by Vocabulary Level, Class, and Treatment

(N = 734-34 = 700)

Criterion Group	Percentages										Percentage of Total
	1	2	3	4	5	6	7	8	Total		
90% Criterion Group	Class	1	2	3	4	5	6	7	8	Total	Percentage of Total
	High	14 (2)	16	13	8	18	3	3	4	79	11
	Middle	9	5	8	10 (2)	8	11	8	4	63	9
80% Criterion Group	Low	5	11 (2)	11 (1)	12 (2)	7 (2)	16 (1)	13 (2)	16 (3)	91	14
	Class	9	10	11	12	13	14	15	16		
	High	2	14	11	12	8 (1)	9	11	5	72	10
80% Criterion Group	Middle	6	8	12	7	11 (2)	14	9	6	73	10
	Low	16	5	5	11 (2)	9 (2)	12	16	15 (2)	90	13

Table 3.9 (Continued)

70% Criterion Group	Class										Percentage of Total
	17	18	19	20	21	22	23	24	Total		
High	4	9	11	12	11	9	3	5	64	9	
Middle	5	16 (1)	8	9	7 (1)	9	4	8	66	9	
Low	12	7	13 (1)	11 (1)	14	14	19 (1)	12 (3)	102 total 700	15	

Numbers within () indicate those students omitted from the data analysis because of excessive absence, late enrollment, or withdrawal from school.

Of the 34 students dropped from the data analysis, nine, or 27%, were in the upper and middle vocabulary groups, while 24, or 73%, were from the low group. These figures point up the fact that frequently corrective procedures cannot operate because the subjects for whom they are intended are not in school where they can benefit from such instruction.

Vocabulary Scores, Grade Equivalent,
and National Percentile Rank

Vocabulary achievement was the concomitant variable used in this study. The mean reading scores for students distributed by treatment, class, and vocabulary level are given in Table 3.10. Tables 3.11, 3.12, and 3.13 give the mean achievement, retention, and attitude toward the unit scores by treatment, class, and vocabulary. The sample population was extremely heterogeneous, with a reading range equivalent of 1.8 for the lowest to 11.9 for the highest group. Grade equivalent ranges by group were: high, 9.4 to 11.9; middle, 6.1 to 7.2; and low, 1.8 to 5.8 as shown in Table 3.14.

The grade equivalent scores translated into national percentile ranks indicate that the mean score of the high vocabulary group was equivalent to the 64 percentile rank, the mean score of the middle vocabulary group was equivalent to the 28 percentile rank, and the mean score of the low vocabulary group was equivalent to the 5 percentile rank (Table 3.14). The majority of students used in this study,

Table 3.10

Mean Reading Scores and Grade Equivalents by Treatment, Class, and Vocabulary Level
(48 items)

Class	1	2	3	4	5	6	7	8	Total	G.E.
90%	H	28.8	31.8	33.5	29.1	30.1	32.3	29.8	30.7	8.2
	M	20.9	21.2	21.0	21.1	20.5	20.4	20.0	20.7	6.5
	L	13.8	12.3	11.6	12.2	12.4	13.0	12.7	12.7	4.2
Class	9	10	11	12	13	14	15	16	Total	G.E.
80%	H	32.5	30.4	30.7	33.6	30.9	31.1	30.8	31.1	8.4
	M	21.3	20.9	21.4	19.6	22.6	20.5	20.6	21.0	6.7
	L	12.6	13.0	14.8	10.8	12.7	13.6	13.4	12.8	4.2
Class	17	18	19	20	21	22	23	24	Total	G.E.
70%	H	29.0	35.1	30.6	33.4	30.3	31.9	32.3	31.4	8.4
	M	21.4	20.3	21.0	20.8	22.1	20.2	22.0	21.2	6.7
	L	12.6	12.6	13.2	12.2	13.7	13.5	13.3	12.8	4.2

Table 3.11
Mean Achievement Scores by Treatment, Class, and Vocabulary Level
(50 test items)

Class	1	2	3	4	5	6	7	8	Total	
90%	H	33.0	38.5	37.8	29.5	30.8	34.3	38.3	35.0	34.7
	M	34.1	33.0	27.1	27.8	16.1	26.7	30.1	22.5	27.2
	L	17.6	17.1	18.3	20.6	19.7	20.1	24.4	21.9	20.0
Class	9	10	11	12	13	14	15	16	Total	
80%	H	36.0	33.2	35.4	39.0	34.3	34.6	30.9	28.2	34.0
	M	29.0	28.9	26.5	23.6	33.0	25.6	21.9	29.2	27.2
	L	19.6	22.6	18.8	17.8	25.0	17.9	20.2	17.1	19.9
Class	17	18	19	20	21	22	23	24	Total	
70%	H	33.5	42.6	34.0	36.8	30.5	30.9	32.3	31.8	34.1
	M	23.6	27.7	28.6	27.8	19.4	21.2	31.8	27.9	26.0
	L	22.5	16.4	21.8	17.3	18.2	18.4	19.7	23.2	19.7

Table 3.12

Mean Retention Scores by Treatment, Class, and Vocabulary Level

(50 test items)

Class	1	2	3	4	5	6	7	8	Total	
90%	H	30.8	35.6	36.2	29.3	29.7	31.3	34.7	32.8	32.9
	M	26.9	25.8	23.5	26.7	16.9	26.0	29.9	21.8	24.7
	L	18.8	15.7	17.3	18.9	17.4	19.3	22.5	21.3	18.9
Class	9	10	11	12	13	14	15	16	Total	
80%	H	35.5	31.9	32.2	37.5	36.6	33.2	30.7	26.2	33.0
	M	24.7	26.3	26.9	23.3	31.6	24.0	21.4	25.8	25.5
	L	18.3	24.6	23.0	18.3	21.6	16.8	18.3	16.1	19.6
Class	17	18	19	20	21	22	23	24	Total	
70%	H	31.8	39.9	32.2	34.4	32.8	30.8	29.7	30.6	32.8
	M	19.5	27.5	27.8	30.4	21.9	18.8	31.0	27.0	25.5
	L	18.6	17.0	21.0	16.3	17.8	18.9	17.8	17.3	18.1

Table 3.13
Mean Attitude Toward the Unit Scores by Treatment, Class, and Vocabulary Level
(17 items)

Class	1	2	3	4	5	6	7	8	Total
90%	H	6.8	6.8	7.8	8.6	7.5	7.4	7.6	7.5
	M	6.3	7.8	7.2	6.5	7.0	7.7	7.2	7.2
	L	7.2	7.4	7.3	7.2	7.0	6.8	7.0	7.1
Class	9	10	11	12	13	14	15	16	Total
80%	H	7.7	7.4	6.6	7.8	7.1	7.8	7.2	7.3
	M	6.9	6.2	7.6	7.8	7.9	6.8	7.9	7.2
	L	7.2	6.3	5.9	6.9	7.5	7.1	7.4	6.9
Class	17	18	19	20	21	22	23	24	Total
70%	H	6.7	7.9	7.4	7.7	6.6	7.6	7.3	7.4
	M	6.3	7.5	8.0	7.4	7.1	6.9	7.9	7.4
	L	7.1	7.1	6.8	7.2	6.9	7.0	7.2	7.1

Table 3.14
Mean Reading Scores, Grade Equivalent, and National
Percentile Rank for Vocabulary Groups

Vocabulary Level	N = 734	% of Sample	Mean Score	Grade Equivalent	Percentile Rank	Grade Equivalent Range
High	218	.30	31.04	8.4	64	7.4 to 11.9
Middle	208	.28	20.98	6.5	28	6.1 to 7.2
Low	308	.42	12.79	4.2	5	1.8 to 5.8

therefore, fell below the national norm for vocabulary as measured by the word meaning section of the Iowa Tests of Basic Skills: Form 6, Level 13.

Unit of Statistical Analysis

The group within the class was selected as the unit of statistical analysis because an important purpose of this study was to determine how students of varying vocabulary ability, especially lower aptitude students, performed under differential criterion levels. Aptitude was measured by a vocabulary test. Three vocabulary groups, high, middle, and low, were determined for each class on the basis of individual scores on the word meaning section of the Iowa Tests of Basic Skills: Form 6, Level 13. These vocabulary groups, three per class, were subsequently employed as the unit of statistical analysis, providing a total of 72 groups for the data analysis.

Statistical Procedures

A 3 x 3 fully crossed analysis of variance (ANOVA) was used with the achievement, attitude, and retention mean scores as criterion measures. This experimental design was used to determine if the class means on each of the three measures differed significantly ($p < .05$) across treatments. The computer program used in the above data analyses was the BMD 12V program (Biomedical Computer Programs, 1973).

Statement of the Statistical Hypotheses

The purpose of this study was to determine the effect of three criterion mastery levels on achievement, attitude, and retention of seventh grade students using a geography instructional unit. To accomplish this purpose, the following statistical hypotheses were tested at the .05 level of significance. The subscript order is class, vocabulary level, treatment, and measure.

Achievement: Main Effects of Treatment

$$H_0: \mu_{..11} = \mu_{..21} = \mu_{..31}$$

This statistical null hypothesis states that there are no statistical differences among achievement means across treatments. This statistical hypothesis was tested against the two-tailed alternative hypothesis that:

$$H_1: \mu_{..11} \neq \mu_{..21} \neq \mu_{..31}$$

This non-directional hypothesis states that there are statistical differences among achievement means across treatments.

Achievement: Main Effects of Vocabulary Level

$$H_0: \mu_{.1.1} = \mu_{.2.1} = \mu_{.3.1}$$

This statistical null hypothesis states that there are no statistical differences among vocabulary groups on the mean posttest scores. This statistical hypothesis was tested against the two-tailed alternative hypothesis that:

$$H_1: \mu_{.1.1} \neq \mu_{.2.1} \neq \mu_{.3.1}$$

This nondirectional hypothesis states that there are statistical differences among vocabulary groups.

Achievement: Interaction

$$H_0: \begin{aligned} &\mu_{.111} - \mu_{.121} - \mu_{.131} = \\ &\mu_{.211} - \mu_{.221} - \mu_{.231} = \\ &\mu_{.311} - \mu_{.321} - \mu_{.331} \end{aligned}$$

This null hypothesis states that on the achievement measure the interaction of treatment and vocabulary levels is a null interaction. This statistical hypothesis was tested against the two-tailed alternative hypothesis that:

$$H_1: \begin{aligned} &\mu_{.111} - \mu_{.121} - \mu_{.131} \neq \\ &\mu_{.211} - \mu_{.221} - \mu_{.231} \neq \\ &\mu_{.311} - \mu_{.321} - \mu_{.331} \end{aligned}$$

This nondirectional hypothesis states that on the achievement measure the interaction of treatment and vocabulary levels is not a null interaction.

The hypotheses for the analysis of variance for each of the remaining effects measures followed the same format. It was not necessary to state each set of hypotheses because of the repetition involved. Therefore, the same hypotheses were applied to the measures of retention and attitude.

Significance Level

This study used the .05 significance level in testing the null hypotheses. This means that a difference as or larger than the obtained one could occur by chance as frequently as five times out of 100. Therefore, the probability

of rejecting a true statistical hypothesis (Type I error) is .05 (Myers, 1966).

A Type II error is the failure to reject a false hypothesis. The relationship between the Type I error and Type II error is inverse. By decreasing the probability of a Type I error, the researcher increases the probability of a Type II error. The selection of a significance level, therefore, reflects a compromise between the relative importance of the two types of errors (Myers, 1966).

The power of a statistical test is defined as $1-\beta$, or the probability of rejecting a statistical hypothesis when it is false and should be rejected. If alpha (Type I error) is held constant, the power of the significance test can be increased by increasing the number of observations in the sample. Therefore, if power is increased, the probability of β (Type II) is decreased (Edwards, 1968). The .05 level was used in this study because it is appropriate to the moderate size of the N utilized, $N = 72$ as indicated in Tables 3.7, 3.10, and 4.1. It represents an adequate balance between the .01 level if the individual ($N = 734$) had served as the unit of analysis or the .10 level if the class ($N = 24$) had been used.

Limitations of the Study

This section discusses three principal limitations of this study that affected its external and internal validity. One limitation, related to the study's external validity, was the use of an available pool of seventh grade students

in 24 classes of the El Paso Independent School District. This population could not be considered representative of the national population. The subjects were below the national average in vocabulary knowledge as measured by the Iowa Tests of Basic Skills: Form 6, Level 13. In addition the ethnic composition of the sample did not follow national ratios. In the present study more than 50% of the students were Mexican American.

A second limitation, also related to the external validity of the study, was the lack of control for the language variable. Those Mexican American students for whom English is a second language were not excluded from the experiment or data analysis, but treated as low ability students on the basis of their scores on the vocabulary test.

The first two limitations described relate to the low aptitude characteristics of the sample, as measured by the concomitant variable. Nevertheless, the sample was appropriate for a study of mastery learning investigating the relationship of different criterion levels and aptitude on achievement, retention, and attitude.

The third limitation of the study, related to the internal validity of the study, was the lack of systematic observations made in the participating classes during the treatment period to ensure that mastery procedures were being followed. Oral directions were provided to all teachers prior to the beginning of the treatment, and each teacher was given written directions, a time schedule to

follow, and a textbook and an answer booklet. The investigator made frequent visits to all participating schools every week. These procedures strengthen the assumption that the teachers and students followed the instructions outlined, but the degree to which individual teachers and students may have deviated from established procedures cannot be determined.

The next chapter will provide the results of the study and a discussion of the findings.

CHAPTER IV

RESULTS AND DISCUSSION OF THE FINDINGS

The present study did not produce evidence supporting the hypothesis that a higher criterion level facilitated learning of materials at the seventh grade level. Analysis of variance was used to test the statistical hypotheses of this study. Geography posttest and delayed posttest achievement and attitude toward the treatment were dependent variables. Vocabulary knowledge, as measured by the Iowa Tests of Basic Skills: Form 6, Level 13 was used as a blocking variable.

The statistical hypothesis that there was no statistically significant difference ($p < .05$) among the treatment groups achieving to the 90%, 80%, or 70% criterion levels was tested at two time intervals: on the 15th day by a posttest after 14 days of instruction and three weeks later by a delayed posttest. An attitude scale was also administered the same day as the posttest.

Presentation of the Findings

Nine hypotheses were tested. The hypotheses were stated in the null form and the probability level for the testing of hypotheses was established at the .05 level of significance. The findings for each of the nine stated hypotheses are

discussed briefly in this section. A more detailed explanation is offered in the second part of this chapter, Discussion of the Findings.

Hypothesis I Treatment and Geography Achievement

H_0 : 1 There is no significant difference in mean achievement across the three treatment levels.

This study sought to determine whether there existed a significant difference in mean posttest achievement among those students who were required to attain a 90%, 80%, or 70% criterion level throughout a learning unit. The computed F ratio to test this null hypothesis was non-significant (Table 4.1). A higher criterion level did not facilitate learning more than a lower criterion level; for example, Table 4.2 shows that mean achievement for the three treatment groups was similar.

Hypothesis II Vocabulary Level and Geography Achievement

H_0 : 2 There is no significant difference in mean achievement across the three vocabulary levels.

This hypothesis states that there will be no significant difference in mean achievement among high, medium, and low vocabulary groups on the posttest geography scores. The computed F ratio for vocabulary effect was significant (Table 4.1). Therefore, the null hypothesis was rejected and the alternative hypothesis was accepted that there are differences in achievement among the three levels of vocabulary. The data analysis indicated a highly significant

Table 4.1
 Analysis of Variance for Treatment,
 Vocabulary Level, and Interaction -
 Achievement Posttest

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares	F Ratio
Total	71	332482.40	4682.85	
Vocabulary Level	2	247246.24	123623.12	92.472*
Treatment	2	563.55	281.77	0.211
Treatment Vocabulary Level	4	449.68	112.42	.084
Replications/ Treatment X Vocabulary	63	84223.04	1336.87	

* Indicates F ratio that is significant at the .001 level..

Table 4.2
Achievement Mean Scores by Treatment and Vocabulary Level

	T ₁	T ₂	T ₃	Row Marginals
	90% Criterion Group	80% Criterion Group	70% Criterion Group	
High Vocabulary Group	$\bar{X} = 34.58$ $s = 3.36$	$\bar{X} = 33.94$ $s = 3.27$	$\bar{X} = 34.04$ $s = 3.97$	34.19
Middle Vocabulary Group	$\bar{X} = 27.19$ $s = 5.80$	$\bar{X} = 27.20$ $s = 3.52$	$\bar{X} = 25.98$ $s = 4.15$	26.79
Low Vocabulary Group	$\bar{X} = 19.95$ $s = 2.39$	$\bar{X} = 19.88$ $s = 2.69$	$\bar{X} = 19.69$ $s = 2.52$	19.84
Column Marginals	27.24	27.01	26.57	

difference at the .001 level. High vocabulary students achieved significantly higher than middle vocabulary students; middle vocabulary students achieved significantly higher than low vocabulary students. Previous knowledge, as measured by vocabulary, was a stronger factor in student achievement than the attainment of differential criterion levels throughout the unit. As indicated in Table 4.2, achievement means across vocabulary levels for the high, medium, and low groups were 34.19; 26.79; and 19.84, respectively.

Hypothesis III Interaction of Treatment and Vocabulary Levels on the Achievement

Measure

H₀: 3 There is no significant interaction between treatment and vocabulary levels.

The computed F ratio for interaction results was non-significant. Vocabulary and treatment did not act together to create a greater effect on student achievement than either of the main effects, treatment and vocabulary, taken into account separately. The interaction effect is the experimental effect created by the combination of treatment and vocabulary over and above any effects associated with treatment and vocabulary considered separately.

Hypothesis IV Treatment and Geography Retention

H₀: 4 There is no significant difference in mean retention across the three treatment levels.

This hypothesis sought to determine whether there existed a significant difference in mean delayed posttest

achievement among those students who were required to attain a 90%, 80%, or 70% criterion level throughout the learning unit. The computed F ratio to test the null hypothesis was non-significant (Table 4.3). The attainment of a higher criterion level did not facilitate retention more than a lower level. As indicated in Table 4.4, the mean achievement score is the same for the 90% criterion level and the 70% criterion level, a finding consistent with the achievement measure noted previously.

Hypothesis V Vocabulary Level and Retention

H₀: 5 There is no significant difference in mean retention across the three vocabulary levels.

This hypothesis states that there will be no significant difference in mean retention among high, medium and low vocabulary groups on the delayed posttest geography scores. The computed F ratio for vocabulary effect was highly significant at the .001 level (Table 4.3). Therefore, the null hypothesis was rejected and the alternate hypothesis was accepted that there are differences in retention among the three vocabulary levels. High vocabulary students achieved significantly higher than middle vocabulary students; middle vocabulary students achieved significantly higher than low vocabulary students. Previous knowledge, as measured by vocabulary, was a stronger factor in student achievement than were differential criterion levels. Table 4.4 indicates the significantly different mean retention scores across vocabulary levels.

Table 4.3
 Analysis of Variance for Treatment,
 Vocabulary Level, and Interaction -
 Achievement Delayed Posttest

Source of Variation	Degrees of Freedom	Mean Squares	Mean Squares	F Ratio
Total	71	299811.84	4222.70	
Vocabulary Level	2	231964.96	115982.48	109.94*
Treatment	2	649.71	324.85	0.30
Treatment X Vocabulary Level	4	754.97	188.74	0.17
Replications/ Treatment X Vocabulary	63	66442.24	1054.63	No Test

* Significant at .001 level.

Table 4.4

Retention Mean Scores by Treatment and Vocabulary Level

	T ₁	T ₂	T ₃	Row Marginals
	90% Criterion Group	80% Criterion Group	70% Criterion Group	
High Vocabulary Group	$\bar{X} = 32.52$ $s = 2.68$	$\bar{X} = 32.96$ $s = 3.64$	$\bar{X} = 32.76$ $s = 3.23$	32.75
Middle Vocabulary Group	$\bar{X} = 24.67$ $s = 3.96$	$\bar{X} = 25.49$ $s = 3.01$	$\bar{X} = 25.47$ $s = 4.77$	25.21
Low Vocabulary Group	$\bar{X} = 18.88$ $s = 2.20$	$\bar{X} = 19.63$ $s = 3.06$	$\bar{X} = 18.07$ $s = 1.45$	18.86
Column Marginals	25.36	26.03	25.43	

Hypothesis VI Interaction of Treatment and
Vocabulary Levels on the Retention Measure

H₀: 6 There is no significant interaction between treatment and vocabulary levels.

The computed F ratio for interaction results was non-significant (Table 4.3). Vocabulary and treatment did not act together to create a greater effect upon student retention than either of the main effects, vocabulary and treatment, taken into account separately.

Hypothesis VII Treatment and Attitude

H₀: 7 There is no significant difference in posttest mean attitude scores across the three treatment levels.

This hypothesis sought to determine whether there existed a significant difference in mean attitude among those students who were required to attain a 90%, 80%, or 70% criterion level throughout the learning unit. The computed F ratio to test the null hypothesis was non-significant (Table 4.5). Attitude toward the unit was not affected by the required criterion level (Table 4.6). The attitude scale represented a continuum ranging from a very positive disposition toward the subject, represented by a value of 10.3, to a very negative disposition toward the subject, represented by a value of 0.6. As indicated in Table 4.6, the three treatment groups sustained a relatively high attitude toward the unit: T₁ and T₃ averaged 7.3, and T₂ averaged 7.1, indicating little, if any, difference among treatment groups.

Table 4.5
Analysis of Variance for Treatment,
Vocabulary Level, and Interaction -
Attitude Posttest

Source of Variation	Degrees of Freedom	Mean Squares	Mean Squares	F Ratio
Total	71	332482.40	4682.85	
Vocabulary Level	2	144.08	72.04	2.95
Treatment	2	33.25	16.62	0.68
Treatment X Vocabulary Level	4	25.66	6.41	0.26
Replications/ Treatment X Vocabulary	63	1539.00	24.42	No Test

Table 4.6
Attitude Mean Scores by Treatment and Vocabulary Level

	T ₁		T ₂		T ₃		Row Marginals
	90% Criterion Group	80% Criterion Group	70% Criterion Group				
High Vocabulary Group	$\bar{X} = 7.5$ $s = .6$	$\bar{X} = 7.3$ $s = .4$	$\bar{X} = 7.4$ $s = .5$	7.4			
Middle Vocabulary Group	$\bar{X} = 7.2$ $s = .5$	$\bar{X} = 7.1$ $s = .7$	$\bar{X} = 7.4$ $s = .6$	7.2			
Low Vocabulary Group	$\bar{X} = 7.1$ $s = .2$	$\bar{X} = 6.9$ $s = .6$	$\bar{X} = 7.1$ $s = .2$	7.1			
Column Marginals	7.3	7.1	7.3				

Hypothesis VIII Vocabulary Level and Attitude

H₀: 8 There is no significant difference in posttest mean attitude scores across the three vocabulary levels.

This hypothesis states that there will be no significant difference in mean attitude among high, medium, or low vocabulary groups on the posttest Attitude Toward Any Subject scale scores. The computed F ratio for the vocabulary effect was non-significant (Table 4.5). Students belonging to the high, middle, or low vocabulary level groups did not differ among themselves on the attitude scale. Therefore, the more stringent 90% criterion level did not result in a less positive attitude toward the unit than the 80% or 70% levels. Table 4.6 shows the small difference among the high, middle, and low vocabulary groups, 7.4; 7.2; and 7.1, respectively.

Hypothesis IX Interaction of Treatment and Vocabulary Level on the Attitude Measure

H₀: 9 There is no significant interaction between treatment and vocabulary levels.

The computed F ratio for interaction results was non-significant (Table 4.5). Vocabulary and Treatment did not act together to create a greater effect on student attitude than either of the main effects, treatment and vocabulary, taken into account separately.

These three univariate tests of significance for the nine hypotheses are summarized in Table 4.7.

Table 4.7

Summary of Univariate Tests of Significance:
Interaction and Main Effects

Statistical Hypotheses (Null)	F	Level of Significance
There are no differences:		
I. Achievement: mean differences for main effects and interaction;		
1. Main Effects: Treatment	0.21	N.S.
2. Main Effects: Vocabulary	92.47	.001
3. Interaction: Treatment by vocabulary	0.08	N.S.
II. Retention: mean differences for main effects and interaction;		
4. Main Effects: Treatment	0.30	N.S.
5. Main Effects: Vocabulary	109.94	.001
6. Interaction: Treatment by vocabulary	0.17	N.S.
III. Attitude: mean differences for main effects and interaction;		
7. Main Effects: Treatment	0.68	N.S.
8. Main Effects: Vocabulary	2.95	N.S.
9. Interaction: Treatment by vocabulary	0.26	N.S.

7

Discussion of the Findings

The results of this study of mastery learning indicated no significant difference by treatment level criterion on the summative measures of achievement, retention, and attitude. Therefore, unlike the finding of Block, higher criterion mastery levels required throughout the unit Population Growth in the United States and Mexico did not produce higher levels of achievement on the summative test. High vocabulary students achieved and retained more of the geography unit than middle or low vocabulary level students. Middle vocabulary level students achieved and retained more of the geography unit than low vocabulary level students (See Table 4.2).

These findings suggest that differences in achievement were mostly a function of the aptitude attributed to individual students at the beginning of instruction. The data are congruent with literature dealing with achievement and individual differences, indicating that treatment is not always a sufficient factor in every attempt to equalize performance (Wright, 1967). Notwithstanding the purported provision of mastery elements of feedback and correction, low aptitude students, as measured by vocabulary knowledge, were unable to overcome their learning difficulties to match higher aptitude students, even when required to meet a high criterion level on formative exercises.

The number of students who attained criterion on the geography summative posttest equivalent to the criterion which had been required on the formative exercises shows that mastery procedures failed to eliminate differences in achievement or retention by vocabulary level. These data are shown in Table 4.8, noting the number of students who attained criterion on the summative test and delayed posttest by treatment and vocabulary level. Table 4.8 also shows that there was little difference, if any, among mean achievement, retention, and attitude scores by treatment. A possible interpretation of this finding is that there was too little difference in the corrective exercises among the criterion levels to make a significant difference in achievement, retention, or attitude. In most of the lessons there was only a difference of one or two items that had to be answered correctly in order to reach criterion. For example, in a lesson requiring 20 responses, the 90% criterion group had to answer 18 of the 20 items correctly to reach criterion and be able to proceed to the next lesson; the 80% group had to answer 16 correctly; and the 70% group had to answer 14 correctly. On a lesson requiring 10 responses, the 90% group had to answer 9; the 80%, 8; and the 70%, 7 in order to meet the prescribed criterion. Appendix B gives the personal score logs for each treatment groups, indicating the minimum number of items to be answered correctly by students to achieve criterion.

Table 4.8

Number of Students Achieving Criterion on the Summative Test by
Treatment and Vocabulary Level and Mean Achievement,
Retention, and Attitude, by Treatment

	90% Criterion Group	80% Criterion Group	70% Criterion Group
Number Attaining Criterion on Geography Posttest by Vocabulary Level	11 (5%) H-8; M-2; L-1	27 (12%) H-23; M-4; L-0	50 (21%) H-35; M-12; L-3
Number Attaining Criterion on the Geography Delayed Posttest by Vocabulary Level	6 (3%) H-3; M-2; L-1	21 (9%) H-18; M-3; L-0	46 (20%) H-31; M-14; L-1
Mean Achievement Score by Treatment	27.24	27.01	26.57
Mean Retention Score by Treatment	25.36	26.03	25.43
Mean Attitude Score by Treatment*	7.3	7.1	7.3

* Attitude scale 0.6, representing an extreme dislike toward the unit, to 10.3,
representing an extreme liking toward the unit.

The significant difference in mean achievement, however, came in terms of vocabulary levels. Most of the students who attained criterion on the geography posttest were high and middle vocabulary level students as noted in Table 4.8. Only four low aptitude students achieved their respective criterion levels. Therefore, feedback and correction procedures failed to eliminate learning problems of lower aptitude students. Retention findings are consistent with the achievement findings. Since the required 90% criterion level did not yield greater posttest achievement results than the attainment of the minimum criterion, there is no reason to assume that a significant difference on the retention measure would result after a three-week delay period for administration of the delayed posttest measure. Therefore, none of the attained performance levels tended to equalize average achievement or retention around a high score.

Student attitude, measured by the Attitude Toward Any Subject scale, was not affected by the differential criterion level. On a scale from 10.3 to 0.6, where 10.3 represents a positive and 0.6 a negative attitude toward the subject, a mean of 7.1 for treatments one and three and 7.3 for treatment two, as shown on Table 4.8, indicates a relatively positive attitude toward the subject. Statistically, the three treatment groups sustained a higher than neutral attitude toward the unit. The attitude data, however, is inconsistent with the observational data on pupil attitude reported by teachers and observed by the investigator in the

pilot study. Observational data confirmed the assumption that the attainment of the more stringent 90% criterion level would result in less positive attitude toward the unit than the 70% or 80% levels. As observed by both teachers and the investigator, the groups that achieved to the 70% or 80% criterion levels showed a more positive attitude toward the treatment because they repeated procedures less. The 90% level was sometimes frustrating to students, especially lower ability students who reportedly showed a desire to quit. Some also attempted to skip the more difficult exercises. Students who failed to attain the 90% criterion level made comments which indicated a dislike for the procedure because the lesson in its entirety was repeated, especially when criterion had almost been attained. Mean attitude scores by vocabulary level as shown in Table 4.6 indicates a small difference among the three levels, even though the higher the vocabulary level, the higher the mean score for that level.

The lack of significant difference may be attributed to two factors. First, the study was a short-term one that introduced new materials and procedures to students who may not have overcome the newness of the treatment within the three week duration of the study. Second, the lack of control to ensure that mastery procedures were strictly adhered to by students may have contributed to a more positive attitude by those who did not, in fact, experience the full impact of the mastery treatment. Overall, however, teachers and the investigator observed that students enjoyed working on the

unit, some taking it home to get ahead. Students often commented positively about the nature of the materials and the fact that they could write in their books and keep them. All of the teachers reported that some low ability students appeared to do much better with the unit than with other types of classwork customarily assigned, even some of those pupils with a limited command of the English language. The frequent reinforcement which these students received was considered by the teachers to be a valuable tool in their learning process.

It has been noted that unit completion time was not statistically reported because of the constraints of a limited time period which did not permit students to complete the treatment unit, thereby defeating mastery learning procedures due to the lack of time. However, data were gathered to indicate the number of lessons completed by treatment and vocabulary level.

More students from the 90% criterion group failed to complete the unit than pupils from either the 80% or 70% criterion groups. While about 20% of the total number of subjects failed to complete the unit, 41% of those assigned to the 90% performance level did not complete the unit. Of those subjects in the 80% and 70% groups, 31% and 28%, respectively, did not complete the unit. However, there was no difference among treatment groups as to the amount of the unit completed. The amount of the unit completed by subjects within each of the three treatment groups was

the same, approximately half of the unit's lessons were finished.

A significant difference did occur, however, in the number of students who did not complete the unit by vocabulary level. A series of t-tests were calculated with the mean number of students not completing the unit by vocabulary level. There was a significant difference at the .05 level between the middle vocabulary group and the low vocabulary group, as well as the high vocabulary group and the low vocabulary group. This finding is consistent with the hypothesis that achieving to high levels of mastery will require additional time. Table 4.9 shows the approximate number and percentage of students not completing the unit and the mean number of lessons completed by treatment and vocabulary level. These data point out the dilemma in school instruction with respect to time needed by lower aptitude students to complete a learning task. If the lower aptitude students who did not complete the unit had been allowed the time they needed to complete it, would their achievement have increased as measured by the posttest?

The findings of this study are contrary to those of Block (1970) who examined both cognitive and affective consequences of four different performance levels--95%, 85%, 75%, and 65%. Even though this is not a comparative study, Block's findings are pointed out because of the effect which the use of differential criterion levels had on measures common to both studies--achievement, retention, and

Table 4.9

Number and Percentage of Students Not Completing the Unit and Mean
Number of Lessons Completed by Treatment and Vocabulary Level

Aptitude by Vocabulary	T ₁				T ₂		
	90% Criterion Group		X̄ Number Lessons Completed	Percent Not Completing Unit	80% Criterion Group		X̄ Number Lessons Completed
	Number Not Completing Unit	Percent Not Completing Unit			Number Not Completing Unit	Percent Not Completing Unit	
High	15	.190	33		3	.042	31
Middle	10	.159	20		11	.151	30
Low	32	.352	23		31	.348	23
Total	57	.40			45	.32	
High Middle Low Total	T ₃				Total	Percent	
	70% Criterion Group			29	25	.114	
	7	.110	21				
	8	.125	23				
	25	.243	88				
	40	.28		142	.100		

140

attitude. Block reported that maximal cognitive learning as measured by achievement, transfer, and retention was produced by the group of eighth graders achieving to the 95% mastery level on a unit of matrix algebra. However, the long run effects on interest and attitude were negative. Maximal interest and attitudes were produced by those attaining the prescribed 85% mastery level. The same group achieved only slightly less than optimal cognitive learning.

The exclusion of 15 subjects from Block's data analysis is important to note. He reports that some students were omitted from the analysis because of their reluctance to participate in the study. The exclusion of such individuals from the data analysis gives spurious results that need careful interpretation. In addition, Block's control group received instruction not representative of the typical classroom. Students in the control group were given the programmed text, told to study it with no further reinforcement, and given the formative test. Control group procedures such as those in Block's study do not appear representative of traditional classroom management where a teacher presumably takes a greater corrective role. Therefore, control treatment procedures should be considered when mastery versus non-mastery conditions are being compared to ensure that control groups are not deliberately assigned inferior instruction, deemed representative of traditional classroom instruction.

Reasons for obtaining non-significant results in this mastery study by criterion level are subject to speculation. The researcher suggests three plausible explanations for the obtained results. First, students learned the mechanics of a mastery procedure but failed to learn the unit content. The overt responses made by students as they proceeded to derive information from a data base in order to complete sentence stems were in themselves insufficient to guarantee mastery of content material. Upon completion of a lesson's sentence stems, it was important that a student study the composed narrative in its entirety to learn it. This important reinforcing procedure was not monitored by the mastery procedures. If a student failed to internalize the learning outcomes, he could not recall them later in the final test. At the end of each of the three parts in the unit text, a 15 item multiple choice test was self-administered using the same procedures employed with the individual lessons. Observational data from teachers and the investigator showed that students who took these three review tests performed poorly and showed considerable surprise at their low scores. This information indicates that students were failing to learn the content as they completed the exercises.

The studies utilizing the FIRM method of instruction support the above conclusion. First, Dale (1972) reported no significant difference in learning between the FIRM and narrative units of instruction, Population Growth in the

United States and Mexico. Pelletti (1973), building upon Dale's study, investigated the role of graphics in geography according to five different modes of presentation. The results indicated that the increased time it took students to complete the FIRM unit, graphics as primary communication with focusing instruments, did not influence achievement. On the contrary, students achieved as well with less time given simpler instructional materials, such as narrative with graphics as reinforcement or narrative only.

A second possible explanation is related to a confounding variable--language. Mastery learning is a strategy aimed primarily at the low achiever. In this study, many of the low achievers were also classified as non-English speakers. Because the materials were written at approximately the seventh grade reading level, predominantly Spanish speaking students probably found the unit more complicated than native English speakers, as observed by the investigator and teachers. While teachers attempted to tutor frequently the Spanish speaking pupils and to supervise closely their progress, the limited amount of time available in a class period prevented the teacher from achieving optimal tutoring conditions.

A third reason for non-significant results is the possible lack of procedural implementation by teachers and students. The teacher role in this study was very important because it called for active participation, that is, constant interaction with students and close supervision

of students. However, it was difficult for the teachers to supervise each child's progress as much as had been initially expected. Students not meeting criterion and failing to repeat the corrective procedure or claiming to have met the criterion, when in fact they had not, could have proceeded through the unit undetected by the classroom teacher for a period of time. In addition, the investigator had changed the class-paced mastery pattern, to an individualized program. The class-paced pattern had originally called for periodic class discussions, supplementary activities or materials, and review procedures to be implemented throughout the treatment at the teacher's discretion. In the present study the emphasis shifted from mastery of material to completing the unit within the time period of 15 days.

The last chapter will provide a summary and conclusions of the study and recommendations for follow-up research.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS FOR FOLLOW-UP RESEARCH

This investigation was an outgrowth of interest in Bloom's construct of mastery learning. Bloom has proposed that under appropriate instructional conditions, virtually all students can learn most of what they are taught. The problem of mastery learning is one of finding a strategy that matches instruction to the student so that he actually spends the time needed to learn. A major premise of mastery learning, therefore, is that diagnostic and prescriptive instruction can reduce, if not completely eliminate, learning difficulties of children customarily attributed to aptitude.

The specific purpose of this study was to investigate the relationship of differential criterion levels and aptitude on student achievement, retention, and attitude. Aptitude was measured in terms of the vocabulary subtest of the Iowa Test of Basic Skills. Vocabulary knowledge was used to establish three vocabulary groups--high, middle, and low. Achievement and retention were measured by the population geography test developed by Dale to accompany the unit Population Growth in the United States and Mexico. Attitude toward the subject by different criterion levels was measured

by the Attitude Toward Any Subject scale of Remmers (short form).

Differential criterion levels, aptitude, and interaction effects generated nine research hypotheses. These may be summarized: there is a significant difference ($p < .05$) by three criterion levels--90%, 80%, and 70%—and by three vocabulary levels--high, middle, and low—and in the interaction effects of criterion and vocabulary on

1. achievement,
2. retention, and
3. attitude.

Procedures

Instruction utilized the text Population Growth in the United States and Mexico (Dale & Rice, 1972) organized according to a method of presentation called the Forced Inferential Response Mode (FIRM). It uses incomplete sentence stems to force students to derive information from a data base to construct a series of sequential responses. The correctly completed stems compose a logical narrative which explain information contained in the data base. In this unit, the data base consisted of maps, graphs, tables, and charts. Each data base provided the content for one lesson.

The mastery procedure consisted of five main steps. First, a student studied a data base as a source of information. Second, he proceeded to complete all of the stems on the corresponding page of the data base. These answers were written by the student. Third, when the stems were completed,

a student obtained an answer booklet from the teacher to check his answers. If a student met the prescribed criterion level, he recorded the score in his personal score log and proceeded to the next lesson. However, if a student failed to attain the prescribed criterion level, he was required to review the original data base and repeat the lesson's exercises on a clean sheet of paper. Fifth, he obtained feedback on the correctness of his responses. A second score was also recorded in the pupil's personal score log. A student then proceeded to the next unit, whether or not the specified criterion had been attained in the correction procedure.

Twenty-four seventh grade classes from the El Paso Independent School District in El Paso, Texas served as the experimental population. From this available pool the classes were randomly assigned to three groups and then treatments were randomly assigned to groups.

An analysis of variance was conducted with each dependent variable--achievement, retention, and attitude--to determine the affect of differential criterion levels on students of varying ability. Vocabulary knowledge was used as a blocking variable. The study was a three-week treatment with a delayed posttest administered three weeks upon completion of the unit text.

Findings

On the geography achievement measure, a higher criterion level did not facilitate learning more than a lower

criterion level. There were highly significant differences in achievement among the high, middle, and low vocabulary level groups. The results were the same on the geography retention measure as on the achievement measure. The attainment of a higher criterion level did not facilitate retention more than a lower level. There was a highly significant difference in retention among the high, middle, and low vocabulary groups. On the attitude measure there was no significant difference in mean attitude across treatment, the 90%, 80%, or 70% criterion groups, nor across vocabulary level, the high, middle, and low groups.

Conclusions

The findings for the main treatment effects were consistent on all four criterion measures--achievement, retention, attitude, and unit completion time. The study did not produce evidence supporting the hypothesis that average achievement would be maximized under the maintenance of more stringent criterion levels. The same lack of relationship between the various levels and a delayed posttest, likewise indicate that more rigid criterion levels did not facilitate a higher degree of achievement. Most important of all was the statistical difference at the .001 level among vocabulary groups on the achievement posttest and delayed posttest, indicating that learning of the geography unit Population Growth in the United States and Mexico was more a function of aptitude, as measured by a word meaning test, than of

differential criterion levels. However, as noted in the chapter on procedures, there is doubt if the pupils actually met the stipulated criterion level. Where students of different aptitude in a class are provided the same treatment, low aptitude students do not have an advantage that does not also affect the top aptitude group. In this study limited feedback-correction procedures did not help the low ability students reach a high level of achievement, irrespective of the criterion level assigned.

Attitude toward the subject was not affected by criterion level. The three criterion groups were almost identical in attitude and indicated a positive attitude toward the treatment. The mean number of lessons completed was not affected by the criterion level. As expected, the higher ability students completed significantly more lessons than the low ability students, as did, also, the middle ability group in comparison with the low ability students. These findings are contrary to those of Block, the only other investigator to examine the effect of various criterion levels on learning. In the Block study the maintenance of high criterion levels throughout a learning sequence did contribute to significantly higher levels of summative achievement. However, Block did not examine the relationship between achievement and aptitude. The findings of this study were consistent with those of other studies employing social studies materials and conducted at the intermediate level (Jones, 1974; Wyckoff, 1974; Fagan, 1975). These studies

have shown that mastery on formative exercises did not contribute to higher levels of achievement by lower aptitude students. In view of these findings, the following observations are made.

Low ability students need special assistance from teachers and the school in overcoming many learning problems. Jones (1974) suggested that mastery procedures coupled with a teacher who is prepared to work closely with the low achievers would appear to offer the disadvantaged student some hope of overcoming learning deficiencies. In this study, the suggestion of Jones to incorporate close teacher supervision and tutoring of low aptitude students was employed in addition to the use of student monitors to assist low achievers. Nevertheless, the results of this study and those of Jones were similar. In view of the fact that self-instructional materials did not work successfully with low aptitude students, even though frequent teacher and student tutoring was provided, a combination of class-paced and individual based mastery materials and procedures is suggested as one means of assisting the slow learner to overcome his learning difficulties. Corrective procedures, in addition to tutoring, include use of alternate learning materials or exercises, small group study sessions, and class discussions.

The use of multiple feedback-correction procedures will require more time for lower ability students than the average and above average achievers. But whether or not the advantage of this practice outweighs the disadvantage is important

to note. First, about 20% of the total number of subjects in this study did not finish the unit. Of this 20%, 41% were from the 90% criterion group, 31% were from the 80% criterion group, and 28% were from the 70% criterion group. Yet, there was no significant difference among the treatment groups on the amount of the unit completed nor the summative achievement performance. The possible advantage of greater achievement in mastery learning approaches must be weighed against the disadvantage of less material covered. In the social studies, where subject matter is much less structured than mathematics, the extra time that mastery procedures appear to require may be unwarranted. The question arises as to whether it is quantity or quality that the school desires most for its students. Today's school administrators have to decide whether the economics of achievement weighed against the additional time that is needed to obtain quality instruction is congruous with their educational goals.

The last section of the chapter will discuss means for correcting the major limitations of this study in a future investigation.

Recommendations for Follow-up

Research

Based upon the findings, observations, and conclusions of the present study, the investigator submits the following specific recommendations for further research relating to the facilitative effects of mastery criterion levels.

This study should be replicated with the recommendations by the investigator that offset the major limitations of the present study. The purpose should remain the same, to study the effect of different criterion levels on the achievement, retention, attitude, and unit completion time of seventh grade students using Population Growth in the United States and Mexico. Aptitude should remain as a concomitant variable, measured by vocabulary, reading ability, or IQ and socioeconomic status.

The first recommendation concerns the duration of the study. The three week duration of the present study was insufficient time to complete the unit in its entirety, 44 lessons, and provide supplementary classroom and group activities to reinforce learning. The study should be extended an additional two to three weeks to enable the unit to be taught under a semi-class-based approach that combines both individual and group based techniques. Because those students who did not complete the unit averaged about 22 completed lessons, the investigator suggests that an additional two to three week extension will be sufficient to enable all students to complete the unit, in addition to enabling the teacher to conduct group based activities as a means of unit reinforcement. This procedure was originally prescribed for this study; however, when the classroom time available was considered along with the length of the unit in the pilot phase of the study, the approach was switched to an almost

completely individualized program to enable most students to complete the unit within a three week period.

In addition to recommending an extended period for the study to provide a semi-class-based strategy, the investigator suggests that alternative corrective procedures be utilized. Besides the corrective alternative which calls for a review of original content material and the use of teacher and student tutoring, two other corrective teaching procedures are suggested. The first is the utilization of small group discussion sessions to review particular lessons or unit sections. Second, the FIRM unit Population Growth in the United States and Mexico has a straight narrative counterpart with no graphics which can be employed as an additional corrective. Students who do not attain criterion can be directed to the same lesson within the narrative text, containing the identical sentence stems in the FIRM unit but in a completed and uninterrupted fashion. A student will study the narrative before repeating lesson exercises of the FIRM unit or taking a researcher-constructed formative test.

The third recommendation is a class-based procedure which should enhance student internalization of content material. Students will be required to explain to the class any given data base in a systematic manner. In explaining a map, for example, the ideal will be for the student to relate systematically the components of a map to one another --the title to the key, the key to the map itself, and the map to the title of the lesson. First, the teacher will

have to demonstrate this technique to the class, using each type of data base in the unit text. A questioning component can be added in which the teacher and the other students raise questions to be answered by the pupil explaining the data base to the class, or other class members as well when the student who is reporting on a data base does not reply to a given question. There are a number of class-based procedures that can be incorporated into the semi-class-paced mastery strategy; pupil led discussions is only one recommended procedure that may help students to better understand the unit Population Growth in the United States and Mexico.

The above recommendations to extend the duration of the study, implement additional correctives, and provide class-based activities and discussions have been made by the investigator to correct the main weaknesses of this study in other related investigations undertaken in the future. Students, especially lower aptitude students, may be more successful in their learning because they will go beyond the mechanics of a mastery procedure to perhaps a greater understanding of the material.

An additional recommendation concerns the need for further mastery learning research in the social studies to ascertain whether hierarchically sequenced subjects such as mathematics lend themselves more readily to mastery procedures than do the social studies in which the sequencing of material may be logical but not necessarily hierarchical.

REFERENCES

REFERENCES

- Ames, L. B. A low intelligence quotient often not recognized as the chief cause of many learning difficulties. Journal of Learning Disabilities, 1, 1968, 45-49.
- Ames, L. B., and Walker, R. N. Prediction of later reading ability from kindergarden Rorschach and IQ scores. Journal of Exceptional Psychology, 55, 1964, 309-313.
- Atkinson, R. C. Computer-based instruction in initial reading. Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton: Educational Testing Service, 1968.
- Ausubel, D. P. Educational psychology: A cognitive view. New York: Holt, Rinehart and Winston, Inc., 1968.
- Bagley, W. C. An essentialist's platform for the advancement of American education. Educational Administration and Supervision, 24 (April 1938), 241-56.
- Barth, J. L., and Shermis, S. S. Defining the social studies: An exploration of three traditions. Social Education, 34 (November 1970), 743-751.
- Bercheck, I., and Tauss, V. Opening classroom: Guiding the transition from traditional to open classroom. Elementary School Journal, 74 (November 1973), 78-83.
- Bereiter, C., and Engelmann, S. Teaching disadvantaged children in the pre-school. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1966.
- Biehler, Robert F. Psychology applied to teaching. Boston: Houghton Mifflin Company, 1971.
- Best, T. On the other side of sesame street. American Education, 10 (May 1974), 6-10.
- Bigge, M. L. Learning theories for teachers. New York: Harper, 1971.

- Block, J. H. The effects of various levels of performance on selected cognitive, affective, and time variables. Unpublished doctoral dissertation, University of Chicago, 1970.
- Block, J. H., ed. Mastery learning: Theory and practice. New York: Holt, Rinehart and Winston, Inc., 1971.
- Block, J. H. Student learning and the setting of mastery performance standards. Educational Horizons, 50 (1972), 183-191.
- Block, J. H. Mastery learning in the classroom: An overview of recent research. Paper presented at the American Educational Research Association, New Orleans, 1973.
- Block, J. H., ed. Schools, society, and mastery learning. Atlanta: Holt, Rinehart and Winston, Inc., 1974.
- Bloom, B. S. Stability and change in human characteristics. New York: John Wiley and Sons, 1964.
- Bloom, B. S. Compensatory education for cultural deprivation. New York: Holt, Rinehart and Winston, 1965.
- Bloom, B. S. Learning for mastery, Evaluation Comment, 1, No. 2 (1968).
- Bloom, B. S. An introduction to mastery learning theory. Schools, Society, and Mastery Learning, ed. by J. H. Block. Atlanta: Holt, Rinehart and Winston, Inc., 1974.
- Bloom, B. S., Hasting, T. J., and Madaus, G. F. Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill Book Company, 1971.
- Bolton, E. B. The Measurement of attitude toward mathematics. Psychological Monographs, 50 (1938), 155-182.
- Brownell, W. A. Criteria of learning in educational research. Journal of Educational Psychology, 39 (1948), 170-182.
- Campbell, Donald T. A review of generalized attitude scales. The Fourth Mental Measurements Yearbook, ed. by O. K. Buros. Highland Park, New Jersey: Gryphon Press, 1953, 46.

- Campbell, Donald T. The Purdue master attitude scales. The Sixth Mental Measurements Yearbook, ed. by O. K. Buros, 1965, 168.
- Carroll, J. B. A model of school learning. Teacher's College Record, 64 (1963), 723-33.
- Carroll, J. B. Problems of measurement related to the concept of learning for mastery. Educational Horizons, 48 (1970), 71-80.
- Clark, Kenneth E. A review of generalized attitude scales. The Fourth Mental Measurements Yearbook, ed. by O. K. Buros. Highland Park, New Jersey. Gryphon Press, 1953, 46.
- Collins, K. M. A strategy for mastery learning in modern mathematics. Unpublished study, Purdue University, Division of Mathematical Sciences, 1970.
- Colon, B. A. A comparison of the performance of seventh-grade students with and without prior knowledge of the objectives of an individualized science program. Unpublished doctoral dissertation, Florida State University, 1970.
- Cook, J. M. Learning and retention by informing students of behavioral objectives and their place in the hierarchical learning sequence. USOE Final Report, 1969.
- Cubberley, E. P. Public education in the United States. Boston: Houghton Mifflin Co., 1934.
- Dale, J. R. An analysis of the effects on achievement using the forced inferential response mode in an intermediate grade population geography unit. Unpublished doctoral dissertation, University of Georgia, 1972.
- Dale, J. R., and Rice, M. J. Population growth in the United States and Mexico. Athens, Georgia: Geography Curriculum Project, University of Georgia, 1972.
- Dalis, G. T. Effect of precise objectives upon student achievement in health education. The Journal of Experimental Education, 39 (1970).
- Duchastel, P. C., and Merrill, P. F. The effects of behavioral objectives on learning: A review of empirical studies. Review of Educational Research, 43 (Winter, 1973), 53-69.

- Duffy, O. B., Clair, T. N., Egeland, B., and Dinello, M. Relationship of intelligence, visual motor skills, and psycho-linguistic abilities with achievement in the 3rd, 4th, and 5th grades: A follow-up study. Journal of Educational Psychology, 63, 1972, 358-362.
- Dumbleton, D. D. The effects of guided inquiry and expository materials on cognitive learning, retention, and transfer in a social studies unit for secondary level students. Unpublished doctoral dissertation, University of Georgia, 1973.
- Dunn, L. M., and Horton, K. B. Peabody language development kits: Manual for Level 2. Circle Pines, Minn.: American Guidance Services, Inc., 1966.
- Ebel, R. Essentials of educational measurement. Englewood Cliffs, N.J.: Prentice-Hall, 1972.
- Educational Policies Commission. The central purpose of American education. Washington, D.C., 1961.
- Edwards, A. L. Experimental design in psychological research. New York: Holt, Rinehart and Winston, Inc., 1968.
- El Paso profile. El Paso, Texas: El Paso Industrial Development Corporation and El Paso Chamber of Commerce, 1974.
- Fagan, J. Transportation and the environment. Athens, Ga.: Geography Curriculum Project, University of Georgia, 1972.
- Fagan, J. The effect of mastery and non-mastery on the achievement and retention of seventh grade students studying a geography unit. Unpublished doctoral dissertation, University of Georgia, in process, 1975.
- Feather, N. T. Effects of prior success and failure on expectations of success and subsequent performance. Journal of Personality and Social Psychology, 3, 1966, 287-298.
- Ferguson, L. W. Personality measurement. New York: McGraw-Hill, 1952.
- Findley, W. G. A comprehensive plan for reorganizing education in these United States now. Educational Horizons, 50, 1972, 167-175.

- Fishburne, R. P. A comparison of programmed and a non-programmed text on evolution for the fifth grade. Unpublished Master's thesis, Georgia Southern College, 1971.
- Fitzpatrick, E. A., ed. St. Ignatius and the ratio studiorum. New York: McGraw-Hill, 1963.
- Flesch, R. How to test readability. New York: Harper, 1949.
- Ford, E., and Walker, V. Public secondary schools of the United States. Washington, D. C.: U.S. Office of Education, 1961.
- Fry, E. Judging readability of books. Teacher Education, (1964), 34-49.
- Gaines, W. G. An application of John B. Carroll's model of school learning to the teaching of anthropology. Unpublished doctoral dissertation, University of Georgia, 1971.
- Glaser, R. Toward a behavioral science base for instructional design. Teaching Machines and Programmed Learning, ed. by R. Glaser. Washington, D. C.: National Educational Association, 1965.
- Glaser, R. Adapting the elementary school curriculum to individual performance. Proceedings of the 1967 Invitational Conference in Testing Problems. Princeton: Educational Testing Service, 1968.
- Gwynn, J., and Chase, J. B. Curriculum principles and social trends. New York: MacMillan, 1969.
- Haynes, C. A. Humanizing instruction in the open classroom. Educational Horizons, 52 (1973), 32-6.
- Healey, J. R., and Stephenson, L. K. Unit mastery learning in an introductory geography course. The Journal of Geography, 74 (January 1975), 7-11.
- Hunt, J. Intelligence and experience. New York: Ronald Press, 1961.
- Jacobs, P., Maier, M., and Stolurow, L. A guide to evaluating self-instructional programs. New York: Holt, Rinehart and Winston, 1966.
- Jarvis, O. T., and Rice, M. J. An introduction to teaching in the elementary school. Dubuque, Iowa: Wm. C. Brown, 1972.

- Johns, Jerry. Children's concepts of reading and their reading achievement. Journal of Reading Behavior, 4 (1972), 56-57.
- Jones, F. G. The effects of mastery and aptitude on learning, retention, and time. Unpublished doctoral dissertation, University of Georgia, 1974.
- Jones, F. G. Functions of Cities. Athens, GA.: Geography Curriculum Project, University of Georgia, 1974.
- Jones, L. B. New patterns of instruction in English composition. Report to the Board of Education. Denver, Colorado, 1962.
- Kapenzi, G. Z. Job training for the disadvantaged. Negro Educational Review, 23 (1972), 29-34.
- Keller, F. S. Goodbye, teacher. . . . Journal of Applied Behavioral Analysis, 1 (1968), 79-89.
- Kersh, B. Y. Programming classroom instruction. Teaching Machines and Programmed Learning, ed. by R. Glaser. Washington, D. C.: National Education Association, 1965.
- Kersh, M. E. A strategy for mastery learning in fifth-grade arithmetic. Unpublished doctoral dissertation, University of Chicago, 1970.
- Kilpatrick, W. H. Foundations of method. New York: Macmillan, 1926.
- Kim, H. Learning rates, aptitudes, and achievements. Unpublished doctoral dissertation, University of Chicago, 1968.
- Kim, H. The mastery learning project in the middle schools. Seoul: Korean Institute for Research in the Behavioral Sciences, 1970.
- Kim, H. Mastery learning in the Korean middle schools. UNESCO Regional Office for Education in Asia, 6, No. 1 (1971), Sec. I, 55-66.
- Kruger, C. G. Social sciences and the humanistic curriculum. Social Studies, 65 (1974), 257-8.
- Lawler, R. M., Dick, W., and Riser, M. Mastery learning and remedial prescriptions in computer-managed instruction. Journal of Experimental Education, 43 (1974).

- Lee, Yung Dug, et al. Interaction improvement studies on the mastery learning project--final report on mastery learning program, April-November, 1971. Educational Research Center, Seoul National University, November, 1971.
- Lindquist, E. F., and Hieronymus, A. N. Iowa Tests of basic skills: Form 6, level 13. Boston: Houghton Mifflin, 1971.
- Lindquist, E. F., and Hieronymus, A. N. Manual for administrators, supervisors, and counselors. Boston: Houghton Mifflin, 1974.
- Lysaught, J. P., and Williams, C. M. A guide to programmed instruction. New York: Wiley, 1963.
- Markle, Susan M. Good frames and bad: A grammar of frame writing. New York: Wiley, 1964.
- Maslow, A. H. Toward a psychology of being. Princeton: Van Nostrand, 1968.
- Merrill, M. D., Barton, K., and Wood, L. E. Specific review in learning a hierarchical imaginary science. Journal of Educational Psychology, 61 (1970), 102-9.
- Merrill, P. F., and Towle, N. J. The effects of the availability of objectives on performance in a computer-managed graduate course. Tech Memo No. 47. Tallahassee: CAI Center, Florida State University, 1972.
- Morrison, H. C. The practice of teaching in the secondary schools. Chicago: University of Chicago Press, 1926.
- Myers, J. L. Fundamentals of experimental design. Boston: Allyn and Bacon, 1966.
- Myers, R. A study of mastery and non-mastery in an introductory college geography course. Unpublished doctoral dissertation, University of Georgia, in process, 1975.
- Oswald, J. M., and Fletcher, J. F. Some measured effect of specificity and cognitive level of explicit instructional objectives upon test performance among eleventh grade social science students. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, 1970.
- Pelletti, J. C. The effects of graphic roles on learning geography materials in the middle grades. Unpublished doctoral dissertation, University of Georgia, 1973.

- Pipe, P. Practical programing. New York: Holt, Rinehart and Winston, 1966.
- Pressy, S. L. A simple apparatus which gives tests and scores--and teaches. School and Society, 23, 1926, 373-376.
- Reed, J. E. An experiment involving use of "English 2600," an automated instructional text. Journal of Educational Research, 55, 1962, 476-84.
- Remmers, H. H. A scale to measure attitude toward any school subject. Lafayette, Ind.: Purdue Research Foundation, 1960.
- Rice, M. J. Variables in mastery learning in elementary social studies. Research proposal submitted to the National Institute of Education, 1973.
- Romberg, A., Shepler, J., and King, I. Mastery learning and retention. Technical Report No. 151, Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, Madison, Wisconsin, 1970.
- Rugg, H., and Shumaker, A. The child centered school. New York: World Book Co., 1928.
- Rushton, E. The Roanoke story. Chicago: Encyclopaedia Britannica Press, 1963.
- Schramm, W. The research on programed instruction: An annotated bibliography. Washington: U.S. Dept. of Health, Education, and Welfare, Office of Education, 1964.
- Shaw, M. E., and Wright, J. M. Scales for the measurement of attitudes. New York: McGraw-Hill, 1967.
- Sheppard, W. C., and MacDermot, H. J. Design and evaluation of a programed course in introductory psychology, Journal of Applied Analysis, 3, 1970, 5-11.
- Sherman, J. G. Application of reinforcement principles to a college course. Paper presented at the annual meeting of the American Educational Research Association, New York, 1967.
- Silance, E. B., and Remmers, H. H. An experimental generalized master scale: A scale to measure attitude toward any school subject. Purdue University Studies in Higher Education, XVI, 1934, 84-88.

Skinner, B. F. The science of learning and the art of teaching. Harvard Educational Review, 24, 1954, 86-97.

Smith, B. O., Stanley, W. O., and Shores, J. H. Fundamentals of curriculum development. New York: Harcourt Brace Jovanovich, 1957.

Smith, M. I. Teaching to specific objectives--second process report (evaluation of test data) for the development of a county-wide articulation in foreign language instruction through common measurement procedures. Stanislaus County (Calif.) School's Office, 1968.

Stoddard, Ann. Impact of integrated history material on attitudes toward self and others of black high school students. Unpublished doctoral dissertation, University of Georgia, 1975.

Stolurrow, L. M. Implications of current research and future trends. Journal of Educational Research, 55, 1962, 519-527.

Strunk, O. Attitude toward psychology as a factor in the judgment of the readability of a psychology textbook. Proceedings from the West Virginia Academy of Science, 29, 1957, 175-179.

Stuempfig, D. W., and Maehr, M. L. The effects of conceptual structure and personal quality of feedback on motivation. Child Development, 41, 1970.

Tatsuoka, M. M. Significance tests univariate and multivariate. Champaign, Illinois: Institute for Personality and Ability Testing, 1971.

Thomas, G. The use of programmed instruction for teaching anthropology in the fifth grade. Unpublished doctoral dissertation, University of Georgia, 1967.

Thorndike, E. L. The principles of teaching, based on psychology. New York: A. G. Seiler, 1906.

Thorndike, E. L. Mental discipline in high school studies. Journal of Educational Psychology, 15, 1925.

Thorndike, E. L. The fundamentals of learning. New York: Teachers College, Columbia University, 1932.

- Traxler, A. E. The relationship between vocabulary and general achievement in the elementary school. Elementary School Journal, 45 (1945).
- Walbesser, H. H., and Carter, H. Some methodological considerations of curriculum education research. Educational Leadership, 26 (1968), 53-64.
- Washburne, C. W. Educational measurement as a key to individualizing instruction and promotions. Journal of Educational Research, 5 (1922), 195-206.
- Washburne, C. W. A living philosophy of education. New York: John Day, 1940.
- Washburne, C. W., Vogel, M., and Gray, W. S. Results of practical experiments in fitting schools to individuals. Supplementary educational monograph, Journal of Educational Research, Bloomington, Ill.: Public School Publishing, 1926.
- Wilhelms, F. T., ed. Evaluation as feedback and guide. 1967 Yearbook, Association for Supervision and Curriculum Development (Washington, D. C.: The Association, 1967).
- Wright, W. Achievement as a function of time: An analysis of selected Stanford achievement test battery results. Unpublished manuscript, University of Chicago, Dept. of Education, 1967.
- Wyckoff, D. B. A study of mastery learning and its effects on achievement of sixth grade social studies students. Unpublished doctoral dissertation, Georgia State University, 1974..

APPENDICES

APPENDIX A

Table of Contents and Three Sample
Lessons from Population Growth
in the United States and Mexico

Appendix A gives one sample lesson from each of the three parts in the unit Population Growth in the United States and Mexico, as shown in the table of contents. The first sample lesson from Part I is "Internal Migration." The second sample lesson from Part II is "Mexican Americans in the United States." The third sample lesson from Part III is "Comparison of Birth Rates and Death Rates in the United States and Mexico." The unit consists of a total of 41 lessons and a 15-item review test for each of the three parts. Students in the study were required to fill in the empty blanks of the sentence stems with a response derived from the study of a lesson's data base. The reference to the unit text follows:

Dale, J. R., and Rice, M. J. Population growth in the United States and Mexico.
Athens, Georgia: Geography Curriculum
Project, University of Georgia, 1972.

TABLE OF CONTENTS

	Page
Instructions	iii
Part 1: CHANGES IN THE MEXICAN POPULATION: 1500-1971	
Introduction	2
1. The Distribution of the Indian Population in North America in 1492	4
2. Food Production and Major Mexican Indian Groups in 1519	6
3. Population Changes in Mexico During the Colonial Period	8
4. A Century of Warfare and Slow Growth: 1820-1920	10
5. Land Lost by Mexico Between 1836-1850	12
6. Population Growth in Mexico Since 1930	14
7. The Decline of Mexico's Birth and Death Rate Since 1930	16
8. Causes for the Decline of the Mexican Birth and Death Rates	18
9. International Migrations to and from Mexico	20
10. Internal Migrations	22
11. Major Urban Areas in Mexico: 1970	24
12. Urbanization in Mexico: 1930-1970	26
13. Recent Occupational Changes in Mexico	28
Review Questions for Part 1	29
Part 2: CHANGES IN THE UNITED STATES POPULATION: 1500-1971	
1. The Distribution and Decline of the United States Indian Population	32
2. European Countries Establish Colonies in the New World	34
3. English Settlements Move Inland: 1625-1780	36
4. United States Population Growth During the Colonial Period	38
5. The Economy and People of the English Colonies	40
6. The United States Census	42
7. The Territorial Growth of the United States: 1790-1860	44
8. Population Changes in the United States: 1790-1860	46
9. Growth of the United States Population: 1790-1860	48
10. Conditions Affecting the United States Immigration Rate: 1790-1860	50
11. Immigration to the United States: 1790-1860	52
12. The Decline of the United States Birth and Death Rates: 1850-1900	54
13. Immigration to the United States: 1860-1900	56
14. Population Growth in the United States: 1900-1971	58
15. The Changing Patterns of Immigration to the United States	60

16. The Changing Source of Immigrants to the United States: 1860-1865	62
17. Mexican Americans in the United States	64
18. Internal Migration in the United States: 1900-1970	66
19. The Migration from Cities to Suburbs	68
20. Long Distance Internal Migration in the United States	70
21. The Growth of the United States Population by States: 1960-1970	72
22. Population Density of the United States	74
Review Questions for Part 2	75
 Part 3: A COMPARISON OF THE POPULATIONS OF THE UNITED STATES AND MEXICO TODAY	
1. A Comparison of Population Growth in the United States and Mexico: 1900-1971	78
2. Comparison of Birth Rates and Death Rates in the United States, and Mexico	80
3. Important Traits of the United States and Mexican Populations	82
4. Racial Groups in the United States and Mexico	84
5. A Model of Population Change	86
6. Estimate of the Populations of the United States and Mexico in the Year 2000	88
Review Questions for Part 3	89
Glossary	92

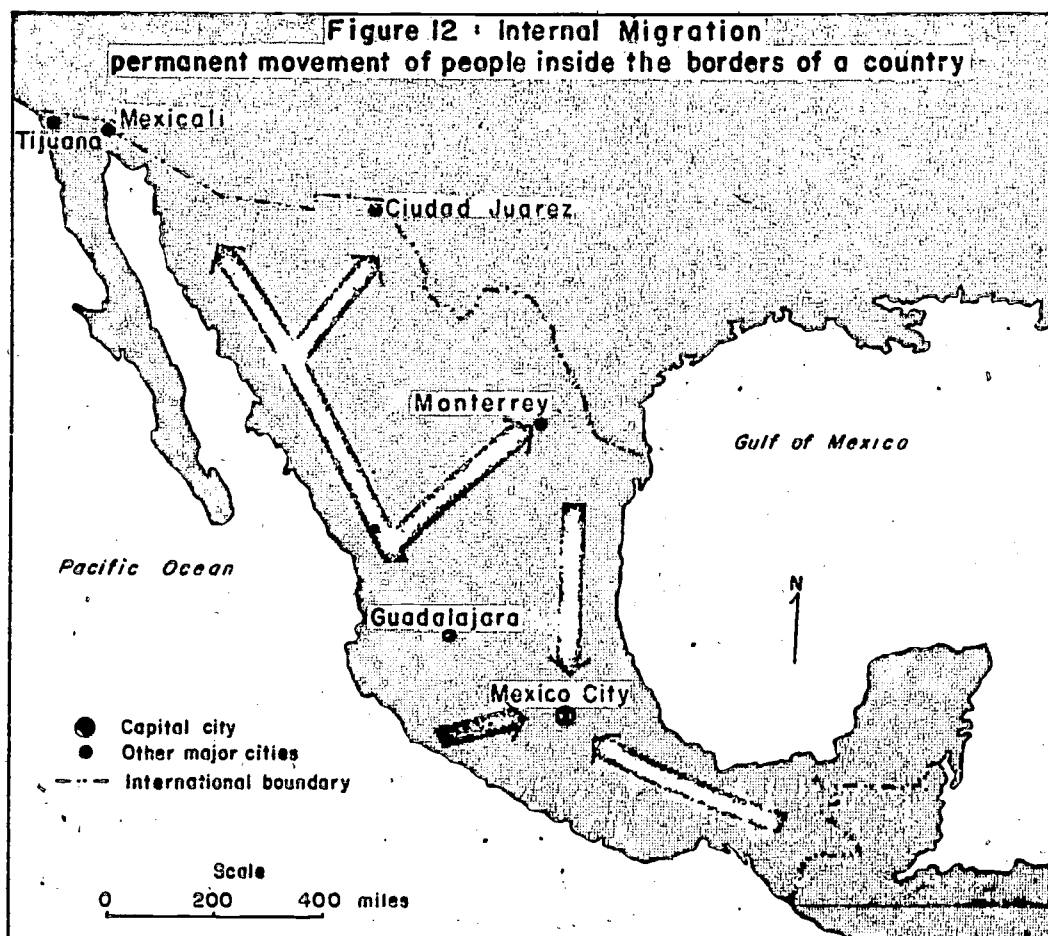


FIGURE 13 CHARACTERISTICS OF RURAL AND URBAN AREAS	
Urban Areas (Cities)	Rural Areas (Farm Areas)
1. Many people living close together. (High population density)	1. Most homes are not close together. (Low population density)
2. Many homes are close together.	2. Most houses have only one family living in them.
3. Many buildings house more than one family.	3. Most people make their living by farming, fishing, or mining.
4. Most people make their living by working in stores, factories, and offices.	

Part 1

10. INTERNAL MIGRATION

1. A second major type of migration involves the permanent movement of people within the borders of a country. This type of migration is called ...

2. In this type of migration the ... of the country are not crossed.

In Mexico today, the most important type of internal migration involves the movement of people from farm areas into cities. Figure 13 shows that:

3. Cities are also called ... areas.

4. In urban areas, the density of population is ...

5. Farm areas are often called ... areas.

6. Places where most of the population makes its living by fishing, farming, or mining are called ... areas.

7. People who live in homes that are very close together usually live in ... areas.

8. The population density of .(1). areas is lower than the population density of .(2). areas.

(1) _____

(2) _____

9. Today many Mexicans are migrating from .(1). areas to live in .(2). areas.

(1) _____

(2) _____

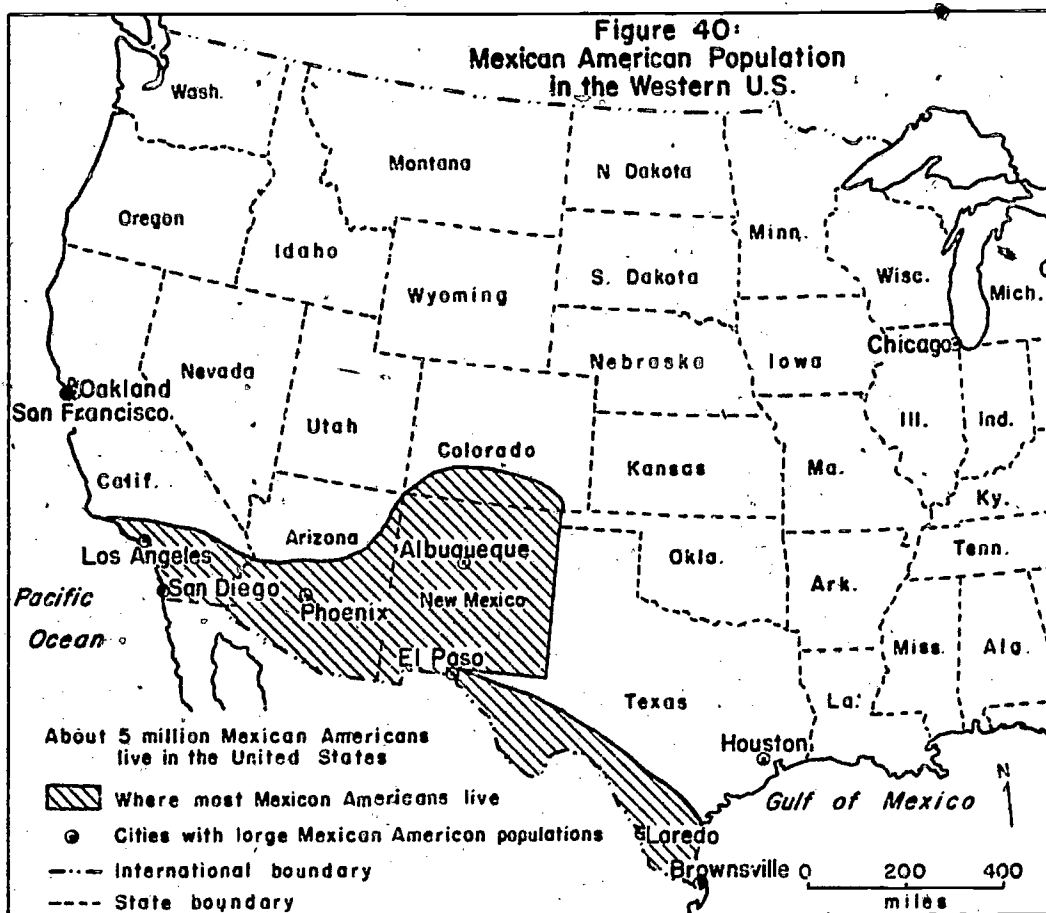
10. One of the major reasons for the rural to urban migration in Mexico is the search for higher paying jobs. The people who move to the city will make their living by working in ... (1)

(2) _____

(3) _____

11. People in rural northern Mexico are pulled northward toward cities along the border of the ...

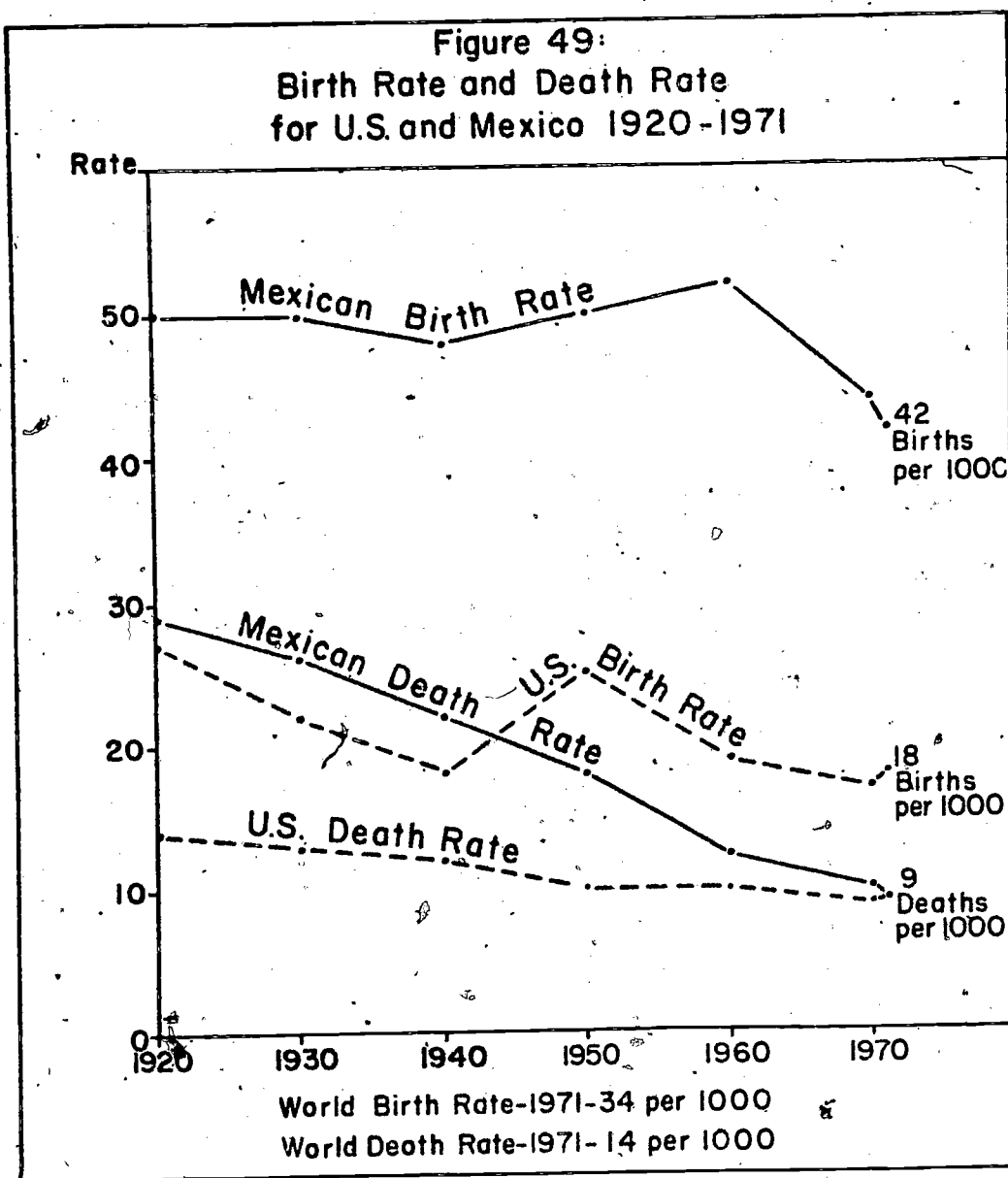
12. People in rural southern Mexico are pulled toward the capital of Mexico, which is ...



Part 2

17. MEXICAN AMERICANS IN THE UNITED STATES

1. Between 1954 and 1964 the nation that sent the largest number of immigrants to the United States was ...
(If you have forgotten, see Figure 39.) _____
2. Mexicans that immigrate to the U.S., or the children of Mexicans that have come at an earlier time, are called _____
3. The total number of Mexican Americans in the U.S. is estimated to be... _____
4. Most Mexican Americans live in the states of ...
(1) _____
(2) _____
(3) _____
(4) _____
(5) _____
5. Since most of these Mexican Americans live in these five states, they make up a very important part of the ... of these states. _____
6. Four of these states are located along the U.S. border with ... _____
7. Outside of this five-state area, the heaviest Mexican American population is found in some cities. Four such cities are ...
(1) _____
(2) _____
(3) _____
(4) _____
8. The number of Mexicans immigrating to the U.S. is larger than the number from any other country. We can therefore expect the Mexican American population in the U.S. to continue to ... _____



Part 3

2. COMPARISON OF BIRTH RATES AND DEATH RATES IN THE UNITED STATES AND MEXICO

1. The last set showed that the total population of the
 (1) _____
 (2) _____
2. It also showed that the Mexican population was growing
 at a ... rate than the United States population. This
 set will try to discover the reasons for the rapid
 growth of the Mexican population.

3. The Mexican birth rate per thousand in 1971 was ...

4. The Mexican death rate for that same year was ... per
 thousand.

5. The Mexican birth rate per thousand was greater than the
 death rate per 1000 in 1971 by ... per thousand.

6. The U. S. birth rate per thousand in 1971 was ...

7. In that same year the U. S. death rate per thousand
 was ...

8. Thus in 1971 the U. S. birth rate was greater than the
 U. S. death rate by only ... per thousand.

9. From these numbers we can see that the major reason for
 the fast growth in Mexico in 1971 was that many more
 Mexicans were born in 1971 than ... in 1971.

10. If we took an average group of 1000 people from the
 Mexican population in 1971, we would find that the
 number of births would be greater than the number of
 deaths by ...

11. If we took a similar group of 1000 people from the U. S.
 population in 1971, we would find the number of births
 would be greater than the number of deaths by ...

12. Thus we can see that Mexico is a country with a high
 (1) _____
 (2) _____
13. The U. S., however, is a country with both a low
 (1) _____
 (2) _____

APPENDIX B

Pupils' Personal Score Logs for the
90%, 80%, and 70% Criterion Levels

PUPIL'S PERSONAL SCORE LOG
(90% Criterion Level)

NAME _____ TEACHER _____ PERIOD _____

SECTION	NO. YOU MUST GET RIGHT	FIRST EXERCISE	SECOND EXERCISE
1.1	15		
1.2	16		
1.3	18		
1.4	19		
1.5	15		
1.6	13		
1.7	13		
1.8	7		
1.9	10		
1.10	14		
1.11	10		
1.12	11		
1.13	7		
REVIEW I	14		
2.1	16		
2.2	9		
2.3	11		
2.4	15		
2.5	14		
2.6	20		
2.7	10		
2.8	7		
2.9	10		
2.10	9		
2.11	10		
2.12	15		
2.13	18		
2.14	20		
2.15	10		
2.16	14		
2.17	14		
2.18	10		
2.19	9		
2.20	11		
2.21	17		
2.22	13		
REVIEW II	14		
3.1	8		
3.2	14		
3.3	14		
3.4	14		
3.5	15		
3.6	6		
REVIEW III	14		

PUPIL'S PERSONAL SCORE LOG
(80% Criterion Level)

NAME _____ TEACHER _____ PERIOD _____

SECTION	NO. YOU MUST GET RIGHT	FIRST EXERCISE	SECOND EXERCISE
1.1	14		
1.2	14		
1.3	16		
1.4	17		
1.5	14		
1.6	11		
1.7	11		
1.8	6		
1.9	9		
1.10	13		
1.11	9		
1.12	10		
1.13	6		
REVIEW I	12		
2.1	14		
2.2	8		
2.3	10		
2.4	14		
2.5	13		
2.6	18		
2.7	9		
2.8	6		
2.9	9		
2.10	8		
2.11	9		
2.12	14		
2.13	16		
2.14	18		
2.15	9		
2.16	12		
2.17	12		
2.18	9		
2.19	8		
2.20	10		
2.21	15		
2.22	11		
REVIEW II	12		
3.1	7		
3.2	13		
3.3	12		
3.4	12		
3.5	14		
3.6	6		
REVIEW III	12		

PUPIL'S PERSONAL SCORE LOG
(70% Criterion Level)

NAME _____ TEACHER _____ PERIOD _____

SECTION	NO. YOU MUST GET RIGHT	FIRST EXERCISE	SECOND EXERCISE
1.1	12		
1.2	13		
1.3	14		
1.4	15		
1.5	12		
1.6	10		
1.7	10		
1.8	6		
1.9	8		
1.10	11		
1.11	8		
1.12	8		
1.13	6		
REVIEW I	11		
2.1	13		
2.2	7		
2.3	8		
2.4	12		
2.5	11		
2.6	15		
2.7	8		
2.8	6		
2.9	8		
2.10	7		
2.11	8		
2.12	12		
2.13	14		
2.14	15		
2.15	8		
2.16	11		
2.17	11		
2.18	8		
2.19	7		
2.20	8		
2.21	13		
2.22	10		
REVIEW II	11		
3.1	6		
3.2	11		
3.3	11		
3.4	11		
3.5	12		
3.6	5		
REVIEW III	11		

APPENDIX C

Treatment Directions to the Student and Teacher

TO THE STUDENT AND THE TEACHER

The unit Population Growth in Mexico and the United States is prepared in the form of a teacher-tutor text.

The population information you are to learn in this book is given in the form of a data base. The data base is presented in the form of a map, a graph, a table, or a chart. Examples of these are:

A Map	page 3; Figure 1
A Graph	page 7, Figure 3
A Chart	page 7, Figure 4
A Table	page 31, Figure 18

On the page opposite the data base is a series of completion statements. The completion statements direct you to study the data base in a particular way. To complete the statement, you must study the map, graph, chart, or table.

Before you begin the unit, your teacher will introduce the different types of data base. Pay careful attention to this introduction so you can get a good idea of how to study a data base. Study especially the title, the key, graph headings, column headings, and other information in the data base.

After your teacher has shown you how to study a data base, then each pupil will begin the unit and proceed at his own pace. Study the data base of a lesson, and then proceed to fill in the blanks on the completion page. Do not try to do it from memory. Look at the data base to find the answer. Then read to yourself the statement with the completed answer. This step of reading your completed statement is important to help you learn the information.

After you have completed an exercise or lesson, take the answer booklet and line it up with the completion exercise. If you get the answers correct put a check mark (✓) on the answer item. If you do not get the correct response, put an (x) mark. Then draw a line through the incorrect response. Add up the number of correct responses (✓) and put the number in your log.

If you got the prescribed number of responses right, or more, you may proceed to the next exercise. If you did not get the passing number right, you are to re-study the data base. If you still do not understand, ask your teacher or a monitor to help you.

After the review, you do the exercise a second time. Take a blank sheet, write your name on it and period, and cover up your answers. Then do the exercise a second time. Check your exercise against your answer book, checking (✓) the correct responses. Enter the number of correct responses on your answer sheet and in your log. Then hand in your answer sheet to the teacher.

APPENDIX D

Questions for Developing Map and Graph
Reading Skills

A TEACHER'S GUIDE TO POPULATION GROWTH IN THE UNITED STATES AND MEXICO: QUESTIONS FOR DEVELOPING MAP AND GRAPH READING SKILLS

FIGURE I:

1. What does Figure I represent? A map? A chart? A graph? (A map)
2. What is the title of the map? (Indian Population Density of North America - 1492)
3. What does the word density, in the title, mean? (Population density refers to how crowded an area is)
4. The title provides four different kinds of information.

Who?	(Indians)
What?	(population density)
Where?	(North America)
When?	(1492)
5. What countries are shown on the map? (Canada, United States, Mexico)
6. Look at the key of the map. How is an international boundary symbolized? (One line, two dots)

Where do you see an international boundary on the map? (Between Canada and the United States, the U. S. and Mexico, and Alaska and Canada)

What does a boundary line on a map signify? (Where one country ends and another begins)
7. What else does the key tell about the map? (The number of Indians per 100 square miles)
8. What does low density mean? (Few or no Indians or 0-10 Indians)
9. What does medium density mean? (Some scattered groups or 11-16 Indians)
10. What does high density mean? (Some large groups or 61-150 Indians)
11. What does highest density mean? (Many large groups or more than 150 Indians)

12. Locate the north arrow on your map (On the east coast of the U. S.). What does the arrow on a map always tell? (North direction)

13. How much distance does the scale on the map represent? (400 miles)

What does the scale mean? (The length of the scale on the map represents 400 miles on actual land surface)

14. Why is it important to study the key of a map? (So that one can understand the information given on the map)

FIGURE 2:

1. What does Figure 2 represent? (A map)

2. What is the title of the map? (Major Indian Groups and Food Production in Mexico - 1519)

Whom does this map tell about? (Major Indian groups)

What is the map about? (food production)

What country is represented? (Mexico)

When was this information true? (1519)

3. What information does the key provide? (International boundary, capital city, and how the Indians produce most of their food)

4. How is an international boundary represented on the map? (One line, two dots)

Do you see a boundary line on the map? (Yes, between Mexico and the U. S.)

5. What does a large dot on the map signify? (The capital city of the Aztecs)

6. How did Indians produce most of their food? (Fishing, farming, hunting, and gathering)

7. How much distance does an inch and a half on the map represent? (400 miles)

8. Find the north arrow (in the Gulf of Mexico). What does the arrow signify? (North direction)

9. What are the names of the major Indian groups shown on the map? (Chichimec, Tarascan, Tarahumar, Aztec, and Mayan)

FIGURE 3:

1. What does Figure 3 represent? (A graph)
2. What is this graph entitled? (Estimate of Mexican Population During the Colonial Period 1519-1810)
3. What was the colonial period of Mexico? (When Mexico was ruled by Spain)
4. How long did Mexico's colonial period last? (From 1519 to 1810 or 291 years)
5. What does the verticle axis, the one going up and down, measure? (Population in millions)
6. What does the horizontal axis, the one going from left to right, measure? (Years)
7. What three groups does the key represent? (Indians, Mestizos, and Whites)
 Which is the largest group? (Indians)
 Which is the second largest group? (Mestizos)
 Which is the smallest group? (Whites)
8. How do you think a person becomes a Mestizo? (Born to parents of mixed races)
9. What does the number 10 on the verticle axis represent? (10 million people)

FIGURE 4:

1. What does Figure 4 represent? (A chart)
2. What is the title of this chart? (The Major Causes for the Decline of the Indian Population: 1519 to 1650)
3. How many years are there between 1519 and 1650? (131 years)
4. What does the word decline mean from using the context and looking at the map? (To decrease in number, to become fewer in number)
5. How many columns does the chart have? (Three)
6. What is the heading of the first column? (Wars)
 What is the heading of the second column? (Diseases)
 What is the heading of the third column? (Forced labor)

7. What does the asterisk after the word "Diseases" tell you? (Look for a footnote)

What does the footnote say? (Diseases killed more Indians than any other cause)

8. Call upon students to read the information in the columns. What were the major types of diseases that caused the death of so many Indians? (Small pox, measles, malaria, and yellow fever)

FIGURE 18:

1. What does Figure 18 represent? (A table)
2. What is the title of this table? (The Decline of the U. S. Indian Population)
3. What is another word for the word "decline" in the title? (Decrease)
4. Read the title again. Whom does this title tell about? (Indians) Where did these Indians live? (The U. S.)
5. How many columns are there in this table? (Seven)
 What is the heading of the first column? (Name of the Region)
 What is the heading of the second column? (Date of First European Contact)
 The third column? (Number of Indians at the Time of Contact)
 The fourth column? (Number of Indians in the Region in 1907)
 The fifth column? (Major Cause of Decline)
 The sixth column? (Major Disease)
 The seventh column? (Largest Tribe in the Region)
6. How many regions are shown on the map or the table? (Ten regions)
7. Locate the Southern Plains region in the first column. Now point to the same region on the map above. What was the largest tribe in that region? (Commanche)
8. How many different characteristics are given to you for each of the ten regions in column one? (Six characteristics beginning with the date of the first European contact and ending with the largest tribe in the region)

9. Is each region listed in the table represented on the map above? (Yes)
10. How are Figure 17 and Figure 18 related to one another? (Figure 17 shows the distribution of those regions given in Figure 18)
11. What was the earliest date of European Contact? (1600) What was the latest date given in Figure 18? (1845) Where did you look to find out the answer to the question? (Looking down the second column)
12. Which of the columns in Figure 18 are totaled at the bottom of the table? (Columns Three and Four) Which total is greater? (Number of Indians at the Time of Contact) What is the difference between these two columns? (583,000) What does this number, 583,000 mean? (583,000 Indians died between 1600 and 1907)

APPENDIX E

Geography Achievement Posttest

Appendix E gives the geography achievement posttest used in this study. The 50-item posttest was developed by J. R. Dale in 1972; the reference to the Dale Study follows:

Dale, J. R. An analysis of the effects on achievement using the forced inferential response made in an intermediate grade population geography unit. Unpublished doctoral dissertation, University of Georgia, 1972.

FINAL TEST FOR UNIT
"POPULATION GROWTH IN THE UNITED STATES AND MEXICO"

DIRECTIONS: SELECT THE ANSWER WHICH BEST COMPLETES THE FOLLOWING SENTENCES. PLACE AN (X) IN THAT SQUARE ON YOUR ANSWER SHEET WHICH CORRESPONDS TO THE ANSWER YOU SELECTED.

1. Before 1500 the density of Indian population in the United States was
 1. far greater than in Mexico.
 2. only slightly greater than in Mexico.
 3. about the same as in Mexico.
 4. less than in Mexico.
2. Most of the Indian groups that lived in southern and central Mexico in 1500 produced their food by
 1. hunting animals like deer and buffalo.
 2. gathering seeds and wild fruits.
 3. farming crops such as maize, beans, and squash.
 4. fishing for trout, salmon, and tuna.
3. The greatest killer of Indians in both Mexico and the United States was the
 1. bullets from the White man's guns.
 2. forcing of Indians to work in gold and silver mines.
 3. diseases brought to the New World from Europe and Africa.
 4. forcing Indians to work on plantations.
4. The major cause for the low rate of immigration to Mexico between 1810 and 1920 was that
 1. Mexico was involved in many foreign and civil wars during this period.
 2. Mexico was already overpopulated in 1810.
 3. diseases made Mexico very unsafe during this period.
 4. the government stopped all immigration to Mexico after 1810.
5. The total population of Mexico in 1971 was about
 1. 12 million.
 2. 52 million.
 3. 72 million.
 4. 102 million.
6. Since 1930 Mexico has gone through a period of
 1. war and rapid population growth.
 2. war and slow population growth.
 3. peace and rapid population growth.
 4. peace and slow population growth.

7. The major reason for Mexico's population growth since 1930 has been
 1. a great decrease in the birth rate.
 2. a great increase in the death rate.
 3. a great decrease in the death rate.
 4. a great increase in the number of immigrants.
8. A migrant who moves out of a country is called an
 1. internal migrant.
 2. internal emigrant.
 3. immigrant.
 4. emigrant.
9. The parts of a country where most of the people make their living by fishing, farming, and mining are called
 1. city areas.
 2. urban areas.
 3. suburban areas.
 4. rural areas.
10. Most of Mexico's major urban places are located
 1. along the shores of the Pacific Ocean.
 2. along the shores of the Gulf of Mexico.
 3. along Mexico's southern border with Guatemala.
 4. at inland sites in the central and southern part of the country.
11. As a larger part of a country's population starts to live and work in urban areas, we can expect the country's
 1. death rate to rise.
 2. birth rate to stay the same.
 3. birth rate to decline.
 4. birth rate to rise.
12. Three European countries that established colonies in North America east of the Mississippi River were
 1. England, Ireland, and Germany.
 2. England, Spain, and Mexico.
 3. England, France, and Ireland.
 4. England, Spain, and France.
13. The major reason the United States government takes a census of the population is to
 1. determine the size of the U.S. population.
 2. determine the amount of taxes each citizen must pay.
 3. decide the number of senators each state will have in the U.S. Senate.
 4. decide the number of representatives each state will have in the House of Representatives.

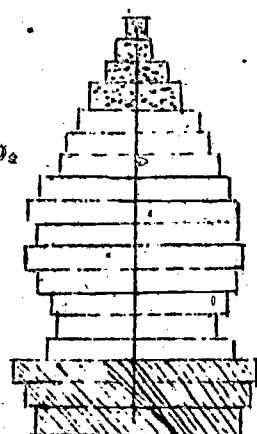
14. The birth rate, death rate, and migration rate are called population
 1. effects.
 2. causes.
 3. variables.
 4. bombs.
15. A shortage of food and a lack of jobs will
 1. pull immigrants into a country.
 2. push emigrants out of a country.
 3. push immigrants into a country.
 4. pull emigrants out of a country.
16. Before 1860 most of the immigrants that came to the United States came from the countries of
 1. England, Ireland, and Germany.
 2. Spain, France, and Italy.
 3. Hungary, Italy, and Poland.
 4. Canada and Mexico.
17. In 1965 most of the immigrants that came to the United States came from
 1. England, Ireland, and Germany.
 2. Mexico and Canada.
 3. Hungary, Italy, and Poland.
 4. England, France, and Spain.
18. The great increase of the United States birth rate after the end of World War II is called the
 1. Great Depression.
 2. Population Explosion.
 3. Baby Boom.
 4. Cold War.
19. Since 1920 the United States
 1. death rate has declined sharply.
 2. birth rate has stayed about the same.
 3. birth rate has declined sharply.
 4. birth rate has moved down, up, and down again.
20. The most important type of United States internal migration that has occurred since 1900 has been
 1. urban to rural.
 2. cities to farms.
 3. rural to urban.
 4. suburban to urban.
21. Since 1940 many Americans have migrated from
 1. cities to suburbs.
 2. cities to farms.
 3. suburbs to cities.
 4. suburbs to farms.

22. The greatest number of Mexican Americans living in United States today live in the
1. southeastern part of the country.
 2. northeastern part of the country.
 3. northwestern part of the country.
 4. southwestern part of the country.
23. The 1970 census showed that
1. all states were not growing at the same rate.
 2. all states were growing at the same rate.
 3. some states were adding large numbers to their populations and none were losing population.
 4. more states were losing population than were gaining population.
24. Since 1910 there has been a strong migration of
1. old people from Florida to New York.
 2. Blacks from the southern cities to the northern cities.
 3. Whites from California to Nebraska and Oklahoma.
 4. Mexican Americans from Ontario to New Mexico.
25. In the United States the area of highest population density runs from
1. Chicago to Los Angeles.
 2. El Paso to Miami.
 3. Boston to Washington.
 4. Washington to Miami.
26. In 1971 the United States had a total population of about
1. 58 million.
 2. 108 million.
 3. 208 million.
 4. 408 million.
27. In 1971 the Mexican population was
1. growing at a slower rate than the United States population.
 2. growing at a faster rate than the United States population.
 3. growing at the same rate as the United States population.
 4. not growing at all.
28. In 1971 the average age of the United States population was
1. older than the average age of the Mexican population.
 2. about the same as the average age of the Mexican population.

3. four times as large as the Mexican population.
 4. eight times as large as the Mexican population.
29. About 10 out of every 100 people in the United States are
1. Mestizos.
 2. Whites.
 3. Blacks.
 4. Mexican Americans.
30. Over 1/2 of the people in Mexico are
1. Spanish.
 2. Mestizos.
 3. Blacks.
 4. Indians.
31. The introduction of modern medicine and improvements in sanitation in many of the underdeveloped nations of the world has caused a rapid
1. increase in the birth rates.
 2. decrease in the birth rates.
 3. increase in the death rates.
 4. decrease in the death rates.
32. By the year 2000 some demographers expect the U.S. population to be about
1. the same size as the Mexican population.
 2. twice as large as the Mexican population.
 3. four times as large as the Mexican population.
 4. eight times as large as the Mexican population.
33. In 1971 Mexico's largest urban place was
1. Mexico City.
 2. Tampico.
 3. Ciudad Juarez.
 4. Tijuana.
34. Some river valleys were good sites of early settlements because they
1. were the homes of only friendly Indians.
 2. had many valuable minerals.
 3. provided good soils and easy transportation.
 4. were free of diseases and insects.
35. Between the years 1790 and 1860 most American families had
1. no children.
 2. about one child.
 3. about three children.
 4. five or more children.

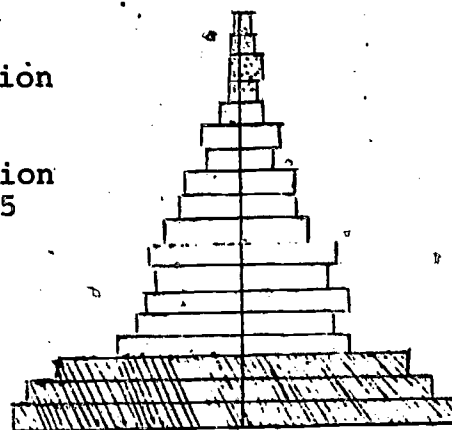
36. The population of the United States in 1971, was
1. smaller than the Mexican population.
 2. about the same size as the Mexican population.
 3. about twice as large as the Mexican population.
 4. about four times as large as the Mexican population.
37. The U. S. Constitution states that a census of the population must be taken every
1. year.
 2. 5 years.
 3. 10 years.
 4. 20 years.
38. The number of women over 65 years old in the United States is
1. less than the number of men over 65.
 2. about the same as the number of men over 65.
 3. far less than the number of men over 65.
 4. greater than the number of men over 65.
39. Many of the countries that are taking their first steps toward modernization have high birth rates and low death rates. This means that they will likely have
1. high growth rates.
 2. low growth rates.
 3. growth rates that are slowing down.
 4. little or no growth.
40. A country with a large part of its population under 15 years old is likely to have
1. high growth rate.
 2. high death rate.
 3. low birth rate.
 4. low growth rate.
41. Many of the more modern nations of the world, like Sweden, Japan, and the United States, have
1. low birth rates and low death rates.
 2. low birth rates and high death rates.
 3. high birth rates and low death rates.
 4. high birth rates and high death rates.
42. A map of population density shows the
1. total population of a country.
 2. number of people living in each state.
 3. average number of people for each square mile.
 4. number of square miles in a country.

Population Pyramid
Country "A"




Males Females

Population Pyramid
Country "B"

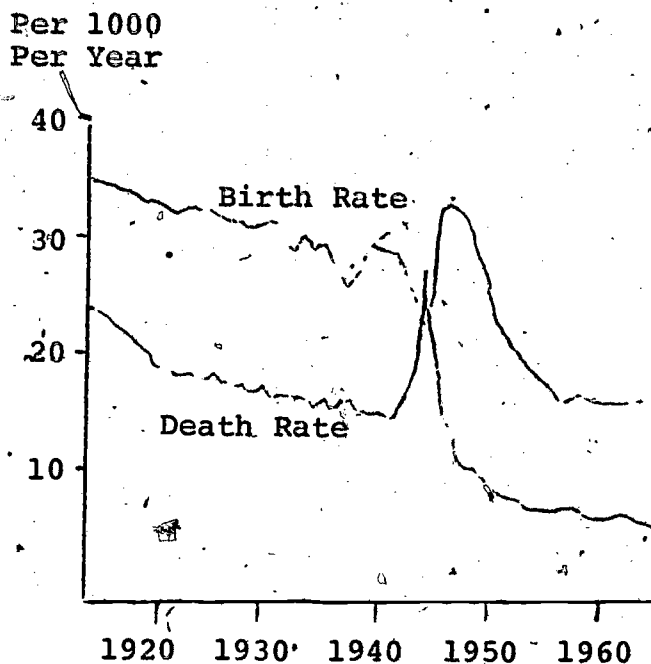


Males Females

 Population
Over 65

 Population
Under 15

43. The part of the population under 15 years old is
 1. greater in Country "A" than in Country "B".
 2. the same in Country "A" as in Country "B".
 3. the same in Country "B" as in Country "A".
 4. less in Country "A" than in Country "B".
44. The birth rate in Country "B" is likely to be
 1. much lower than in Country "A".
 2. only slightly lower than in Country "A".
 3. about the same as in Country "A".
 4. much higher than in Country "A".
45. The part of the population over 65 years old is
 1. greater in Country "A" than in Country "B".
 2. the same in Country "A" as in Country "B".
 3. the same in Country "B" as in Country "A".
 4. less in Country "A" than in Country "B".
46. The government of Country "B" would have to be more concerned with
 1. building old age homes.
 2. providing hospitals for old people.
 3. building more schools.
 4. building homes for newly weds.



47. Since 1920, the birth and death rate in this country has
1. stayed the same.
 2. dropped slowly.
 3. increased slowly.
 4. increased very rapidly.
48. One explanation for the change in the birth and death rate of this country between 1940 and 1950 could be
1. a time of peace.
 2. good working conditions.
 3. war and hard times.
 4. a great flow of immigrants coming in.
49. Population growth in this country would have been the greatest in
1. 1929.
 2. 1939.
 3. 1949.
 4. 1959.
50. The birth rate and death rate of this country in 1960 would be most like that of
1. the United States in 1790.
 2. Mexico in 1810.
 3. Mexico in 1930.
 4. the United States in 1970.

APPENDIX F

Attitude Toward Any Subject

Scale, Short A Form

Appendix F gives the Attitude Toward Any Subject Scale, short A form, that is derived from the original long form developed by Silance and Remmers in 1934. The reference to the long form scale follows:

Silance, E. B., & Remmers, H. H. An experimental generalized master scale: A scale to measure attitude toward any school subject.

The reference to the shortened form developed by Remmers is:

Remmers, H. H. A scale to measure attitude toward any school subject. Lafayette, Indiana: Purdue Research Foundation, 1960.

NAME _____

TEACHER _____

SCHOOL _____

PERIOD _____

ATTITUDE TOWARD GEOGRAPHY UNIT

DIRECTIONS: Please read each of the following statements carefully. Put a check mark () if you agree with the statement. Put a cross (X) if you disagree with the statement. If you simply cannot decide about a statement, you may place a question mark (?) beside it.

1. ___ This unit is profitable to everybody who takes it.
2. ___ No matter what happens, this unit always comes first.
3. ___ I would not advise anyone to take this unit.
4. ___ My parents never had this unit, so I see no merit in it.
5. ___ I am not interested in this unit.
6. ___ All lessons and all methods used in this unit are clear and definite.
7. ___ This unit is a good unit.
8. ___ This unit reminds me of Shakespeare's play--"Much Ado (to do) About Nothing."
9. ___ I look forward to this unit with horror.
10. ___ This unit has an irresistible attraction for me.
11. ___ This unit is a good pastime.
12. ___ This unit will benefit only the brighter students.
13. ___ I don't believe this unit will do anybody any harm.
14. ___ I haven't any definite like or dislike for this unit.
15. ___ This unit is a waste of time.
16. ___ I am willing to spend my time studying this unit.
17. ___ Any student who takes this unit is bound to be benefited.

APPENDIX G

Bibliography of Mastery Learning Studies
at the College Level

BIBLIOGRAPHY

- Airasian, P. W. An application of a modified version of John Carroll's model of school learning. Unpublished Master's thesis, University of Chicago, 1967.
- Austin, S. M. and Gilbert, K. E. Student performance in a Keller plan course in introductory electricity and magnetism. Mimeographed manuscript, Michigan State University, Cyclotron Laboratory, undated.
- Behr, M. J. A study of interactions between structure-of-intellect factors and two methods of presenting concepts of modulus seven arithmetic. Unpublished doctoral dissertation, Florida State University, 1967.
- Biehler, R. F. A first attempt at a learning for mastery approach. Educational Psychologist, 7 (1970), 7-9.
- Born, D. G. Instruction manual for the development of a personality instruction. Center to Improve Learning and Instruction, University of Utah, 1971.
- Born, D. G. and Herbert, E. W. A further study of personalized instruction for students in a large university class. Journal of Experimental Psychology, 40 (1971).
- Collins, K. M. A strategy for mastery learning in freshman mathematics. Unpublished study, Purdue University, Division of Mathematical Sciences, 1969.
- Corey, J. R., McMichael, J. S., and Tremont, P. J. Long-term effects of personalized instruction in an introductory psychology course. Paper presented at the annual meeting of the Eastern Psychological Association, Atlantic City, New Jersey, 1972.
- Corey, J. R., Valente, R. G., and Shamow, N. K. The retention of material learned in a personalized introductory psychology course. Mimeographed manuscript, C. W. Post College of Long Island University, 1971.
- Ferster, C. Individual instruction in a large introductory course. The Psychological Record, 18 (1968), 521-532.

- Fletcher, O. J., and Tyler, D. E. Teaching veterinary pathology by autotutorial and learning for mastery methods. Journal of the American Veterinary Medicine Association, 161 (1972), 65-70.
- Findley, W. G. A comprehensive plan for reorganizing education in these United States now. Educational Horizons, 50 (1972), 167-175.
- Gallup, H. F. The introductory psychology course at Lafayette College, Easton, Pennsylvania: A description and tentative evaluation. Paper presented at the Midwest Psychological Association Meeting, April, 1969.
- Gentile, J. R. A mastery strategy for introductory educational psychology. Unpublished materials, State University of New York at Buffalo, Department of Educational Psychology, 1970.
- Green, Ben A. A self-paced course in freshman physics. Unpublished paper, Massachusetts Institute of Technology, Education Research Center, 1969.
- Hapkiewicz, W. G. Mastery learning options in teaching educational psychology. Mimeographed manuscript, Michigan State University, 1971.
- Harrisberger, Lee. Self-paced individually prescribed instruction. Engineering Education, 61 (1971), 508-510.
- Healey, J. R., and Stephenson, L. K. Unit mastery learning in an introductory geography course. The Journal of Geography, 74 (January 1975), 25-31.
- Hoberock, L. L. Personalized instruction in mechanical engineering. Engineering Education, 61 (1971), 506-7.
- Keller, F. S. Goodbye, teacher. . . . Journal of Applied Behavioral Analysis, 1 (1968), 79-89.
- Koen, B. V. Self-paced instruction for engineering students. Engineering Education, 60 (1970), 735-736.
- Lawler, R. M., Dick, W., and Riser, M. Mastery learning and remedial prescriptions in computer-managed instruction. Journal of Experimental Education, 43 (1974), 45.
- Lloyd, K. E., and Knutzen, N. A self-paced programed undergraduate course in the experimental analysis of behavior. Journal of Applied Behavioral Analysis, 2 (1969), 125-133.

- Mayo, S. T., Hunt, R. C., and Tremmel, F. A mastery approach to the evaluation of learning statistics. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, Illinois, 1968.
- Merrill, M. D., Barton, K., and Wood, L. E. Specific review in learning a hierarchical imaginary science. Journal of Educational Psychology, 61 (1970), 102-109.
- Moore, J. W., Mahan, J. M., and Ritts, C. A. An evaluation of the continuous progress concept of instruction with university students. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois, 1968.
- Myers, R. A study of mastery and non-mastery in an introductory geography course. Unpublished doctoral dissertation, University of Georgia, 1975 (in process).
- Myers, W. A. Operant learning principles applied to teaching introductory statistics. Journal of Applied Behavioral Analysis, 3 (1970), 887-892.
- Semb, G., Hopkins, B. L., and Hursh, D. E. The effects of study questions and grades on student test performance in a college course. Mimeographed manuscript, Department of Human Development, University of Kansas, 1972.
- Sheppard, W. C., and MacDermot, H. J. Design and evaluation of a programmed course in introductory psychology. Journal of Applied Analysis, 3 (1970), 5-11.
- Sherman, J. G. Application of reinforcement principles to a college course. Paper presented at the annual meeting of the American Educational Research Association, New York, 1967.
- Sjogren, D. D. Achievement as a function of study time. American Educational Research Journal, 4 (1967), 337-344.
- Postlewait, S. N., Novak, J. D., and Murray, H. An integrated experience approach to learning with emphasis on independent study. Minneapolis: Burgess Publishing Company, 1964.
- Tierney, M. L. A comparison of feedback/correction procedures used in mastery learning strategies. Unpublished Master's thesis, University of California, Santa Barbara, 1973.

Welser, J. R., Lamar, C. H., and Wilsman, N. J. An exemplary program in comparative anatomy utilizing learning for mastery minicourses and multi-media instruction. British Journal of Medical Association (1972).

APPENDIX H

Request for Authorization to
Conduct Research in the
El Paso Public Schools

EL PASO PUBLIC SCHOOLS
Department of Research

REQUEST FOR AUTHORIZATION TO CONDUCT RESEARCH IN THE EL PASO PUBLIC SCHOOLS

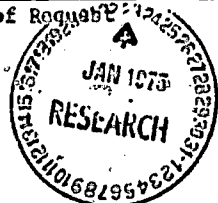
All requests from students and faculty personnel at the University of Texas at El Paso should be submitted to the Office of the Dean, School of Education, at the University. Other requests should be sent to the Director of Research and Evaluation, El Paso Public Schools

1. Name of person initiating request Gloria Contreras
Telephone No. 404-542-7265 Address 107 Dudley Hall, U. of Ga., Athens, Ga. 30602
Identify Instructor, Course Title, and No. if research is University related
A dissertation in partial fulfillment of requirements for an Ed.D. degree.
Dr. Marlen J. Rice, major professor, University of Georgia
2. Name(s) of individual(s) expected to be personally involved in the conduct of the research Gloria Contreras, doctoral candidate Position or Title Research assistant
3. Title of Research Project The Effects of Various Criterion Mastery Levels on Achievement, Attitude, Retention, and Learning Time of Seventh Grade Students
4. Nature of the project. (Give a brief outline of the purpose, method, and general plan of investigation on not more than two pages. Attach to this form.)
5. Collection of data. (Describe specifically the procedures to be used in the collection of data on not more than two pages. Attach to this form.) Give beginning and ending dates of this procedure.
6. Test questionnaire or other instruments to be used. (Give name; attach sample, and give dates when you expect to be in each school where the instrument is to be used.)
7. Schools or Departments from which information is to be collected (List by name if known) Students Involved (Indicate specifically how subjects will be selected) Teachers or other school employees

	Grade Level	Number	Grade Level	Number
See attached listing.				
8. Date and method whereby findings will be made available to the El Paso Public Schools A copy of the dissertation will be provided by August 1975.

1/20/75

Date of Request



AUTHORIZED:

Gloria Contreras
Signature of 1 (above)

Director of Research

Date:

8/74

4. Nature of the Project

THE EFFECTS OF VARIOUS CRITERION MASTERY
LEVELS ON ACHIEVEMENT, ATTITUDE, RETENTION
AND LEARNING TIME OF SEVENTH GRADE STUDENTS

Purpose. The purpose of this study is to determine the effects of three criterion mastery levels (90, 80, and 70 per cent) on achievement, attitude, retention, and time needed to learn by seventh grade students using a geography unit of instruction.

Mastery learning, whether group based or self-paced, is a teaching strategy utilizing a sequential design whereby each segment, containing specific instructional objectives, is to be mastered before instruction begins on the next segment. Operationally, a student must achieve every segment of the unit to a pre-set mastery performance level. This level represents a proportion of correct responses, that is, 90 per cent, 80 per cent, and 70 per cent correct. A feedback/correction component and the allocation of sufficient time to study should theoretically enable most students to achieve criterion mastery performance.

Treatment Materials. Treatment materials for this study will consist of a FIRM geography unit, Population Growth in the United States and Mexico. FIRM is an acronym for the Forced Inferential Response Mode, a self-instructional text which uses incomplete sentence stems to force a student to derive information from a data base in order to construct a series of sequential responses. When correctly completed, the stems and the responses compose a logical narrative which interprets information contained in the data base, that is, a map, a chart, a graph, or a table.

The unit is divided into three parts. Part I contains 16 different data base figures, Part II contains 30 different data base figures, and Part III contains 7 data base figures. Sentence stems from each data base figure

must be answered correctly to the 90, 80, or 70 per cent criterion levels before a student can proceed to the next data base figure. The review tests at the end of each part must also be achieved to a given criterion level before proceeding to the next part.

Failure to achieve a prescribed criterion level will require a feedback learning corrective whereby the teacher directs a student to the section within a segment where learning deficiencies occur. Other corrective procedures will include small group discussions, tutoring by peers who have attained criterion mastery, or teacher-student tutoring.

A final evaluation will be administered to all students upon completion of the unit as an end-of-sequence appraisal of unit objectives. The final summative test and three review tests at the end of each major part will be teacher scored; other segments will be student scored with general teacher supervision.

Teacher Preparation. Teachers will meet with the researcher on February 27, 1975 from 8:30 to 11:30 a. m. for purposes of learning 1) basic concepts of demography, 2) basic knowledge concerning population change in Mexico and the United States, 3) basic skills in map and graph analysis, and 4) basic procedures for a mastery learning strategy.

Teacher participants will be provided with a set of materials Population Growth in the United States and Mexico. They will be demonstrated a model and sequence of map-graph analysis and then take turns practicing individually a replication of the model.

An alternative to the proposed in-service workshop will consist of the researcher meeting with each teacher on an individual basis. Written instructions will be provided, discussed, and clarified.

5. Collection of Data

Duration of the Study. The proposed study is a short-term study that will extend over a period of approximately three weeks. A vocabulary test will be given prior to treatment in order to determine two vocabulary levels--above grade level and below grade level. Approximately three weeks will be allowed for the treatment, including time required to administer the formative tests and employ the correction/feedback procedures of a mastery strategy. One day will be used to administer the summative posttest and attitude scale. A delayed posttest will be administered three weeks after the treatment ends. The delayed posttest will measure the degree of retention evidenced since the posttest. (See next page, chart for data collection schedule)

DATA COLLECTION

Tentative Date for Administration	Procedures	Rationale
1. Feb. 27, 1975	Teachers receive materials and written instructions for mastery learning procedure	Experimental design requires that instructions to teachers be uniform
2. Feb. 27, 1975	Randomly assign classes to treatment groups	Requirement of experimental design
3. Feb. 24 or 25 or 26 (at teachers' convenience)	Teachers administer 17 minute vocabulary section of the Iowa Tests of Basic Skills: Forms 5 and 6	Knowledge of word meaning correlates highly both with ability to read and to achieve in school subjects. By blocking on the knowledge of word meaning variable, to determine high, medium, and low ability groups, the researcher can observe the efficiency of the mastery procedures with students of varying aptitude, especially low aptitude students.
4. March 3, 1975 to March 20, 1975 (classes may finish unit sooner)	Classes begin unit of instruction	
5. March 21, 1975	Summative test and attitude scale are administered. Administration time is one class period.	To determine the effects on achievement and attitude of students achieving to the 70, 80, and 90 per cent criterion levels
6. April 11, 1975	Summative test re-administered. Administration time is one class period.	To test the effects of retention of students achieving to the 70, 80, and 90 per cent criterion levels

7.

School	Teacher	Grade Level	No. of Classes	Approximate No. of Students
Crosby	Aune	7	6	180
White	Gomersall	7	7	210
Guillen	Payne	7	2	60
Canyon Hills	Rope	7	3	90
Henderson	Duncan	7	<u>6</u>	<u>180</u>
			24	720

APPENDIX I

Time Table for Pilot and Experimental
Study

TIME TABLE FOR PILOT AND EXPERIMENTAL
STUDY

Date	Procedure
February 2, 1975	Pilot classes randomly assigned to treatment groups.
February 3, 1975	Pilot treatment began: Instructions given to pupils and 17 minute vocabulary test administered.
February 4, 1975	Pilot classes began unit of instruction.
February 21, 1975	Summative test and attitude scale administered to pilot classes. Teachers received unit text and answer booklet.
February 27, 1975	Orientation meeting with teachers participating in the final study and classes randomly assigned to treatment.
March 3, 1975	Treatment began: instructions given to pupils and 17 minute vocabulary test administered.
March 21, 1975	Summative test and attitude scale administered.
April 11, 1975	Delayed posttest administered.